

University of Alberta Library



0 1620 2908740 8

University of Alberta

Multi-scale Statistical Models of Shape and Appearance: Application to
Modeling Variations in Anatomical Structures

by

Rami Zewail

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biomedical Engineering

Department of Electrical and Computer Engineering

©Rami Zewail

Spring, 2010

Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Mathematical Foundations of Quantum Mechanics
Abstract: A rigorous treatment of the mathematical foundations of quantum mechanics, focusing on the Hilbert space formalism and the spectral theory of self-adjoint operators.

Author: [Name]

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Department of Mathematics

Edmonton, Alberta

Supervisor: [Name]

Abstract
[Name]
[Year]

This thesis is a contribution to the mathematical foundations of quantum mechanics. It is organized into four main parts. The first part is devoted to the review of the basic concepts of quantum mechanics, including the Hilbert space formalism, the spectral theory of self-adjoint operators, and the theory of distributions. The second part is devoted to the study of the scattering theory, including the definition of the scattering operator, the asymptotic completeness, and the theory of resonances. The third part is devoted to the study of the quantum field theory, including the definition of the quantum fields, the Feynman diagrams, and the renormalization theory. The fourth part is devoted to the study of the quantum gravity, including the definition of the quantum metric, the Einstein equations, and the theory of black holes.

The results of this thesis are presented in the form of a series of papers, which are available in the form of a book. The book is published by the University of Alberta Press. The book is available in both English and French. The book is available in both print and electronic form. The book is available in both hardcover and paperback form. The book is available in both English and French. The book is available in both print and electronic form. The book is available in both hardcover and paperback form.

Abstract

Examining Committee

Nelson Durdle, Electrical and Computer Engineering, University of Alberta

Hong Zhao, Electrical and Computer Engineering, University of Alberta

Mrinal Mandal, Electrical and Computer Engineering, University of Alberta

Marek Reformat, Electrical and Computer Engineering, University of Alberta

Sergiy Vorobyov, Electrical and Computer Engineering, University of Alberta

Nilanjan Ray, Computing Science, University of Alberta

Farida Cheriet, Polytechnique Montreal

Abstract

This thesis deals with challenging problems related to the interpretation and quantification of medical images. We develop and apply novel statistical models of shape and appearance to improve the support of various medical practices. We are particularly concerned with two spine-related tasks: the segmentation of vertebral structures in x-ray images, and the analysis of vertebral deformities. These tasks are relevant for various clinical activities such as the diagnosis of spine-related diseases, surgical planning, and studies of disease progression. Within this context, this thesis makes contributions in the following areas:

First, in an attempt to overcome the challenges associated with the modeling of high-dimensional shape space, we propose a novel multi-scale statistical shape model. The new model uses concepts of sparsity, best basis selection and Independent Component Analysis (ICA) to extract multi-scale modes of deformations. This method allows for the construction of a localized non-linear shape model with clinically-relevant modes of deformation.

Second, we present a novel framework for shape analysis and classification of vertebral structures by using pathology-related Wavelet-ICA shape parameters. Within a statistical framework, we are able to establish a statistically relevant relationship between the wavelet-ICA modes of deformations and clinical information. The framework is then extended to achieve multi-vertebrae classification by using Wavelet-ICA parameters. Thanks to the locality of the features, the new classification scheme can assess the condition of several vertebrae simultaneously. Our results show that significant differences are found between groups of healthy and pathological vertebrae.

Third, we develop a novel method for vertebral segmentation by using localized appearance models and a multi-scale shape prior. The new segmentation algorithm

incorporates the following: (i) a method for salient point detection in the non sub-sampled Contourlet domain, (ii) Contourlet-based local appearance profiles, and (iii) a multi-scale shape prior to drive the segmentation process. We evaluate the ability of the proposed method to accurately segment normal and pathological vertebral structures.

Finally, we tackle the challenges related to the accurate modeling of the texture and scarcity of training data. We present a new sparse texture appearance model based upon the Contourlet Transform and Independent Component Analysis (ICA). The new model benefits from the non-linear approximation power of Contourlets to achieve high reduction rates. ICA is then employed to capture localized texture variations. We also describe a general framework that integrates our developed shape and texture models in a unified framework for synthesis of photo-realistic pathological x-ray images.

Acknowledgments

I am first and above all eternally thankful to God Almighty, for it was with his will and guidance that I have completed this work, and he has bestowed infinite other blessings on me throughout my life. I also wish to express my deep thanks and gratitude to the following people and institutions.

My supervisor, Dr. Nelson Durdle, for his guidance and support during the course of my PhD program. Also, I am grateful to Dr. Hong Zhao for serving as a co-supervisor for this thesis.

All my professors, for the major roles they played in my overall education. I would like to especially thank Dr. Dil Joseph and Dr. Mrinal Mandal from the U. of Alberta. Also, I am greatly indebted to Dr. Magdy Saeb for the role he played during my undergraduate and Master study.

My colleagues at the Medical Imaging Lab - Ahmed ElSafi, Lino Rameriz, and Peter Ajemba for their support and encouragement.

All the wonderful people and friends whom I was honoured to meet over the past five years.

The US National Library of Medicine for providing me with the spine x-ray images dataset. Also, thanks goes to Dr. Thomas Deserno at Aachen University of Technology, Germany, for providing me with the database for the medical image retrieval experiments.

The Informatics Circle of Research Excellence (iCORE), and the Killam Trusts for their financial support of this work.

All the administrative staff in the Electrical and Computer Engineering department at the U. of Alberta.

Last, but not least, I am deeply indebted and thankful to my family for its continuous support and encouragement.

Table of Contents

1. INTRODUCTION.....	1
1.1 Motivation.....	1
1.2 Medical Context.....	2
1.2.1 Anatomy of human spine.....	3
1.2.2 Common vertebral pathology	3
1.3 Challenges and Contributions	6
1.3.1 Challenges	6
1.3.2 Contributions	8
1.4 Thesis Organization	11
2. BACKGROUND AND RELATED WORK	12
2.1 Introduction.....	12
2.2 Shape representation	13
2.2.1.1 Spline-based representation.....	14
2.2.1.2 Fourier Descriptor	15
2.2.2.1 Landmark-based representation	15
2.2.2.2 Medial axis-based representation	16
2.2.2.3 Implicit Level-Set representation	17
2.3 Deformable Models.....	17
2.3.1 Active Contours (Snakes)	18
2.3.2 Physical models	20
2.3.3 Statistical shape models.....	21
2.4 Active Shape Models	21
2.4.1 Point Distribution Model	22
2.4.2 Local appearance model	23
2.4.3 Model search procedure.....	24
2.5 Active Appearance Model	25
2.5.1 Shape model	26
2.5.2 Texture model	26
2.5.3 Image Search	28
2.6 Related Work	28
2.6.1 Statistical Shape Modeling	29
2.6.2 Statistical Appearance Modeling	30
2.6.3 Model-based segmentation methods.....	32
2.6.4 Segmentation and analysis of vertebral structures.....	33

3. MULTI-SCALE MEDICAL IMAGE ANALYSIS TOOLS: WAVELETS AND BEYOND.....	36
3.1 Introduction.....	36
3.2 Wavelet transform.....	37
3.2.1 Definition.....	37
3.2.2 Wavelets in medical image analysis	39
3.3 Gabor Transform.....	43
3.3.1 Definition.....	43
3.3.2 Gabor transform in medical image analysis	46
3.4 Steerable Pyramid	47
3.4.1 Definition.....	47
3.4.2 Steerable pyramids in medical image analysis	49
3.5 Contourlet Transform.....	50
3.5.1 Definition.....	50
3.5.2 Contourlets in medical image analysis	53
3.6 Non Subsampled Contourlet transform (NSCT).....	54
3.6.1 Definition.....	54
3.6.2 Non Subsampled Contourlet Transform in medical image analysis.....	56
4. MULTI-SCALE SHAPE MODEL USING BEST BASIS SELECTION FRAMEWORK AND INDEPENDENT COMPONENT ANALYSIS	58
4.1 Introduction.....	58
4.2 Limitations of PCA-based statistical modeling.....	59
4.3 Data Description	61
4.4 Shape modeling as a basis selection problem	62
4.4.1 Multi-scale analysis	62
4.4.2 Sparsity	63
4.4.3 Statistical independence.....	63
4.5 Methodology	64
4.5.1 Multi-scale shape representation	67
4.5.1.1 Wavelet Packets	67
4.5.1.2 Optimal representation using best basis selection	69
4.5.2 Constructing the shape model.....	74
4.5.2.1 Spectral clustering	74
4.5.2.2 Statistical modeling using Independent Component Analysis (ICA).....	75
4.6 Experiments	78
4.6.1 Modes of deformations	79
4.6.2 Shape estimation.....	81
4.6.3 Robustness to noise	85

4.7 Discussion	86
4.8 Conclusion	87

5 VERTEBRAL SHAPE ANALYSIS AND CLASSIFICATION USING LOCALIZED, PATHOLOGY-RELATED SHAPE PARAMETERS..... 88

5.1 Introduction	88
5.2 Methodology	89
5.2.1 Statistical analysis of vertebral structures by using Wavelet-ICA representation.....	89
5.2.1.1 Establishing clinical relevance by using regression analysis.....	90
5.2.1.2 Discriminant Analysis.....	93
5.2.2 Multi-vertebrae classification	93
5.2.2.1 Multi-scale Wavelet-ICA shape representation	95
5.2.2.2 Feature selection	95
5.2.2.3 Shape-based classification.....	97
5.3 Experiments and Results	99
5.3.1 Shape prediction in presence of pathology	101
5.3.2 Evaluation of Wavelet-ICA shape analysis framework.....	103
5.3.3 Evaluation of the multi-vertebrae classification scheme	105
5.3.2 Effect of feature selection on classification accuracy.....	111
5.4 Discussion	112
5.5 Conclusion	113

6. PRIOR-DRIVEN SEGMENTATION USING LOCAL APPEARANCE PROFILES IN CONTOURLET DOMAIN AND MULTI-SCALE SHAPE MODEL 115

6.1. Introduction.....	115
6.2. Contributions.....	117
6.3 Multi-scale salient point detection in Contourlet domain	118
6.3.1 Background.....	118
6.3.2 Methodology.....	119
6.3.2.1 Salient point detector.....	120
6.3.2.2 Incorporation of local appearance knowledge.....	126
6.4 Vertebrae segmentation using contourlet-based salient point matching and multi-scale shape prior.....	133
6.4.1 Methodology.....	133
6.4.1.1 Construction of local appearance profiles	134
6.4.1.2 Construction of the multi-scale shape prior.....	138
6.4.1.3 Segmentation and model search procedure	138
6.4.1.4 An alternative approach: Contourlet-based livewire best-path search	144

6.4.2 Comments	146
6.5 Content-based medical image retrieval using Contourlet-based salient point detector	148
6.5.1 Background.....	148
6.5.2 Methodology.....	149
6.5.3 Comments.....	151
6.6 Experiments and Results.....	152
6.6.1 Evaluation of the NSCT-based salient point detector.....	152
6.6.1.1 Localization.....	154
6.6.1.2 Robustness to rotation	155
6.6.1.3 Robustness to noise	156
6.6.1.4 Robustness to change in illumination.....	158
6.6.1.5 Robustness to non-rigid transformations.....	159
6.6.1.6 Appearance-based salient point detector	160
6.6.2 Evaluation of contourlet-based local descriptor	161
6.6.2.1 Robustness to noise	162
6.6.2.2 Robustness to change of view	162
6.6.3 Evaluation of segmentation algorithm.....	164
6.6.4 Evaluation of the proposed medical image retrieval method.....	168
6.7 Conclusions.....	173
7. SIMULATION OF PATHOLOGICAL DEFORMATIONS USING LOCALIZED MODELING OF SHAPE AND TEXTURE VARIABILITY	174
7.1 Introduction.....	174
7.2 Sparse texture appearance model using Contourlet-ICA	175
7.2.1 Contourlet decomposition.....	176
7.2.2 Non-linear approximation.....	178
7.2.3 ICA modelling in compressed Contourlet domain	179
7.3 Image synthesis framework	181
7.4 Experiments and Results.....	187
7.4.1 Evaluation of the Contourlet enhanced texture appearance model.....	187
7.4.2 Evaluation of the image synthesis framework	193
7.5 Conclusions.....	195
8 SUMMARY AND FUTURE WORK.....	197
8.1 Summary	197
8.2 Future research directions	199
8.3 Final Statement	201

BIBLIOGRAPHY	202
APPENDIX A	219
APPENDIX B	221

List of Tables

Table 3.1 Properties of some well known wavelet families	40
Table 4.1 Total energy captured for a number of target compression percentages, with the size of training set kept fixed to 25.....	73
Table 5.1 Correct classification rates of the proposed classification scheme, and PCA-based classification scheme for detection of anterior osteophyte in lumbar vertebrae.....	107
Table 5.2 Area under curve curves for ROC curves of the proposed classification scheme and the PCA-based classification scheme.....	111
Table 5.3 Specefity at various sensitivities for the Lumbar vertebrae L1-L5.....	111
Table 5.4 Effect of various feature selection methods on classification accuracy.....	112
Table 6.1 Evaluation of robustness of different salient point detectors to non rigid deformations of the object of interest.....	160
Table 6.2 Evaluation of the robustness to noise for various local descriptors.....	162
Table 6.3 Segmentation accuracy of proposed method per vertebra	166
Table 7.1 Performance of Contourlets versus wavelets in compression of spine texture space .	188
Table 7.2 Performance of Contourlets versus wavelets in compression of face texture images	188

List of Figures

Figure 1.1 Anatomy of spine. (a) Lateral view of spine segments: Cervical, Thoracic, Lumbar, Sacrum, and Coccyx, (b) Lateral anatomy of a vertebra [1]	3
Figure 1.2 Lateral Radiographs for human spine with various pathologies (a) Lumbar spine with mild fracture in L3, (b) Lumbar spine with wedge fracture in L3, and crush in L2, (c) spine vertebrae with various fractures, (d) Lumbar spine with osteophytes in L1,L2,L3, (e) Zoom-in showing disc space narrowing example	4
Figure 1.3 The Genant Semi-quantitative grading system [4]	5
Figure 2.1 Example of 2D shape under different similarity transformations.	13
Figure 2.2 Parametric shape representation using (a) polygonal approximation, (b) B-Spline approximation	14
Figure 2.3 Outline of brain structure by using landmark-based representation (from [21])	16
Figure 2.4 (a) Simulated corpus callosum image, (b) medial-axis representation (From [22])	17
Figure 2.5 Implicit Level set representation. A curve is given by a set of points onto the level set function Φ	17
Figure 2.6 Classification of deformable models	18
Figure 2.7 Example of snake-based segmentation (a) Intensity CT image of heart left ventricle (b) edge detected image (c) initial contour (d-f)deformable contour moving toward left ventricle boundary (from [29])	20
Figure 2.8 Modes of vibrations of Finite Element model for (a) synthetic square shape, (b) shape of lumbar spine	21
Figure 2.9 First two modes of variations of PDM model for hand shapes	23
Figure 2.10 First two modes of appearance model of the lumbar spine	27
Figure 2.11 General framework for the work presented in the thesis.....	28
Figure 3.1 Model of the human low-level vision system [95]	36
Figure 3.2 Similarity between receptive fields in mammalian primary visual cortex and Gabor wavelets [96]	36
Figure 3.3 Wavelet decomposition of a two-dimensional signal.....	39
Figure 3.4 Image decomposition by using XY separable 2D wavelet filters.....	39
Figure 3.5 Wavelet decomposition of an x-ray image of the lumbar spine	40
Figure 3.6 Information shift property for four-level orthogonal wavelet decomposition	41

Figure 3.7 Illustration of the limitations of wavelets in representing contours. (a) Wavelet representation of a contour, (b) an x-ray image of lumbar spine with noisy contours, (c) wavelet-based edge map	42
Figure 3.8 (a) One-Dimensional Gaussian function in time and frequency domain, (b) real and imaginary parts of Gaussian function	44
Figure 3.9 2-D Gabor filter (a) Real part; (b) Imaginary part.....	45
Figure 3.10 Gabor Filters. (a) Filter parameters in frequency domain; (b) bank of filters in frequency domain [102]......	46
Figure 3.11 Analysis of lumbar spine x-ray by using Gabor Filterbank (six orientations at a fixed scale)	47
Figure 3.12 First level of Steerable pyramid decomposition.....	48
Figure 3.13 Illustration of the spectral decomposition performed by a Steerable pyramid	49
Figure 3.14 Analysis of lumbar spine x-ray by using Steerable pyramid.....	50
Figure 3.15 Contourlet decomposition by using ideal Pyramidal Directional Filterbank. The horizontal and vertical axes refer to the horizontal and vertical frequencies of an image. The coloured region shows the frequency components present in the image. The white region refers to the frequency band filtered out.	51
Figure 3.16 Division of frequency spectrum by using Contourlet transform.(a) Subspaces generated by LP, (b) Subspaces generated by DFB, (c) Complete division generated by PDFB	53
Figure 3.17 Example of basis functions of contourlet transform.....	54
Figure 3.18 Non Sub-sampled contourlet transform. (a) NSCT filterbank. (b) Idealized frequency partitioning, (c) NSCT basis functions.....	55
Figure 3.19 Non subsampled pyramid decomposition. Frequency response of the building block of the non subsampled pyramid.....	56
Figure 3.20 Frequency response of the building block of the Non Subsampled Directional Filterbank.	56
Figure 4.1 Examples of x-ray images of the human lumbar spine	58
Figure 4.2 Two vertebrae with anterior osteophytes. (a) Lumbar vertebra with upper left-hand corner osteophyte, (b) lumbar vertebra with upper and lower left-hand corner osteophyte..	59
Figure 4.3 First mode of variation of a PCA-based model of lumbar spine	60
Figure 4.4 Examples for failure of PCA-based model to capture localized vertebrae deformations	61
Figure 4.5 Building training dataset for shapes of lumbar spine. (a) Sample x-ray image of lumbar spine, (b) Boundary points selected by an expert, (c) Interpolated shape of lumbar spine	62
Figure 4.6 Shape alignment using Procrustes method: (a) unaligned shapes of lumbar spine, (b) aligned shapes of lumbar spine.	65

Figure 4.7 Block diagram of the proposed multi-scale Wavelet-ICA shape model	66
Figure 4.8 The tree of wavelet packet decomposition. Each doublet (k,i) represents the depth(k), and the number (i) of the nodes.	68
Figure 4.9 Example of shape decomposition using wavelet packets. (a) Mean shape, (b) Mean shape plus variations in low frequency bands, (c) Mean shape plus variations from all frequency bands.	69
(the dotted shape is the original shape).....	69
Figure 4.10 Probability distribution of the multi-scale shape representation of the shapes in the training set (Kurtosis = 90.1136)	71
Figure 4.11 An example for shape reconstruction using the proposed multi-scale representation. 74 (a) original shape, (b) reconstructed shape from the optimal multi-scale space (50% reduction). . 74	
Figure 4.12 Illustration of the spectral clustering algorithm - undirected graph representation with nodes (n1, n2, n3, n4), and weighted edges	75
Figure 4.13 Construction of multi-scale independent Wavelet-ICA basis (WICA)	78
Figure 4.14 Examples of x-ray images with outlined vertebrae contours	79
Figure 4.15 Evaluation of the statistical independence of the basis functions for PCA-based model, ICA-based, and the proposed BB-ICA model.....	80
Figure 4.16 Multi-scale modes of the deformations of the proposed multiscale shape model.	81
Figure 4.17 Average reconstruction error for the following shape models: PCA-based, best-basis PCA (BB-PCA), and the proposed multi-scale best-basis ICA model (BB-ICA)	83
Figure 4.18 Reconstruction of a ground truth test shape with models constructed by using 10 training samples: (a) GT, (b) PCA-based model, (c) BB-ICA model	83
Figure 4.19 Zoom-in to the reconstruction of two L2 lumbar vertebrae by using BB-ICA, and PCA-based models.....	84
Figure 4.21 Reconstruction from a noisy test shape by using 20 training samples. (a) Ground truth noisy shape, (b) PCA-based reconstruction, (c) BB-ICA based reconstruction.....	85
Figure 4.22 Effect of noise on the reconstruction error of the proposed model.	85
Figure 5.1 Vertebral Shape Analysis framework by using Wavelet-ICA representation	90
Figure 5.2 Block diagram for the proposed vertebral classification scheme.	94
Figure 5.3 Example of a WICA-shape parameter that is spatially localized to L4 lumbar vertebra.	97
Figure 5.4 Lumbar spine x-ray image from the National Library of medicine dataset [108].	100
Figure 5.5 Vertebrae with anterior osteophytes. (a) Label points selected by expert radiologist. Points 8, 9 characterize osteophyte pathology, (b) two lumbar vertebrae with the 9-points model (b) Contour of Lumbar vertebra with upper left-hand corner osteophytes, (c) Contour lumbar vertebra with upper and lower left-hand corner osteophytes.	101

Figure 5.6 Local modes of the deformations for L4 vertebra	102
Figure 5.7 Shape prediction for normal and abnormal vertebrae by using PCA-based model (first row), and the proposed model (second row). (a,d) are normal vertebrae, (b,c,e,f) are vertebrae with osteophytes pathology.....	102
Figure 5.8 Regression coefficients from the investigation of the relationship between the localized deformation modes and the osteophyte pathology for lumbar vertebrae L1-L3.Statistical significance levels of the localized deformation modes are indicated by the color code: black ($p<0.001$), dark gray ($p<0.01$), and light gray ($p<0.05$)	104
Figure 5.9 Regression coefficients from the investigation of the relationship between the localized deformation modes and the osteophyte pathology for lumbar vertebrae L4-L5.Statistical significance levels of the localized deformation modes are indicated by the color code: black ($p<0.001$), dark gray ($p<0.01$), and light gray ($p<0.05$).....	105
Figure 5.10 ROC curves for detection of anterior osteophytes in L1 vertebra	108
Figure 5.11 ROC curves for detection of anterior osteophytes in L2 vertebra	109
Figure 5.12 ROC curves for detection of anterior osteophytes in L3 vertebra	109
Figure 5.13 ROC curves for detection of anterior osteophytes in L4 vertebra	110
Figure 5.14 ROC curves for detection of anterior osteophytes in L5 vertebra	110
Figure 6.1 General framework for knowledge-driven methods presented in this chapter.....	118
Figure 6.2 Flowchart of the proposed multi-scale contourlet-based salient point detector	121
Figure 6.3 NSCT basis functions for two scales with different angular resolutions at each scale	123
Figure 6.4 Example x-ray image and corresponding Contourlet-based feature strength image. ...	125
Figure 6.5 Example of ground truth salient points that define the lumbar spine in the x-ray images	127
Figure 6.6 Extraction of Contourlet-based local descriptor for a salient point.	129
Figure 6.7 Building blocks of the Gaussian Mixture Model	130
Figure 6.8 Proposed vertebrae segmentation algorithm. (a) Model learning framework (b) Segmentation framework	134
Figure 6.9 Examples of training x-ray images with outlined contours and characteristic salient points.....	135
Figure 6.10 NSCT image decomposition at different scales and orientations (NSCT Feature Images).	136
Figure 6.11 Model search procedure using contourlet-based local appearance profiles and multi-scale shape prior.....	139
Figure 6.12 Contourlet-based salient point detection. (a) Input image, (b) strength image, (c) input image with extracted salient points.....	140

Figure 6.13 Illustration of TPS warping procedure. (a) Two sets of points used to estimate the TPS parameters, (b) estimated TPS warp, (c) Warped output shape	143
Figure 6.14 Content-based Image retrieval using Non-subsampled contourlet features	149
Figure 6.15 Images used for evaluation of the NSCT-based salient point detector. (a) Synthetic image,.....	152
(b) VanGogh Image, (c) X-ray image of lumbar spine.....	152
Figure 6.16 Localization of various salient point detectors. (a) Harris detector, (b) Laplacian detector, ...	154
(c) Phase Congruency detector (PC-based), and (d) NSCT-based detector.....	154
Figure 6.17 Localization of various salient point detectors in medical images. (a) NSCT-based detector, (b) Harris detector, and (c) Phase Congruency detector.....	155
Figure 6.18 VanGogh test image at various rotation angles.....	156
Figure 6.19 Robustness to rotation of various salient point detectors	156
Figure 6.20 VanGogh test image with added Gaussian noise : (a) clean image, (b) 0.01 noise variance, (c) 0.06 noise variance.....	157
Figure 6.21 Effect of variations in noise levels on the repeatability rate for two different salient point detectors.....	157
Figure 6.22 VanGogh test image with various levels of brightness	158
Figure 6.23 Effect of brightness level on the repeatability rate for various salient point detectors	159
Figure 6.24 Two x-ray images where the object of interest is related by non-rigid deformations	160
Figure 6.25 Example for salient point detection using NSCT-based method (a) with no prior appearance knowledge, (b) with prior appearance knowledge	161
Figure 6.26 Test images with changing viewing angle	163
Figure 6.27 Robustness of various salient point detectors to change in viewing angle.....	163
Figure 6.28 Examples of x-ray images of the lumbar spine used in evaluation of the segmentation algorithm.....	164
Figure 6.29 Segmentation results for three different images by using: Active Shape model (ASM) (a,b,c), ...	165
and proposed segmentation method (d,e,f).....	165
Figure 6.30 Effect of using the proposed Multi-scale prior on segmentation accuracy (a) No prior is used, (b) localized multiscale prior is employed	166
Figure 6.31 Comparison for the delineation capabilities of the proposed contourlet-livewire method (b, d, f), and the traditional livewire algorithm (a, c, e).	167
Figure 6.32 Examples of the IRMA medical image database	169
Figure 6.33 Example of intra-class variations for images in the IRMA database	170

Figure 6.34 Average retrieval rates versus number of retrieved images for the following CBIR methods: Gabor-based salient points (Gabor-local), NSCT-based salient points (NSCT-local), and NSCT-based global response (NSCT-global).	171
Figure 7.1 Block diagram for the Contourlet enhanced texture appearance model (CTICA-AM).....	176
Figure 7.2 (a) Face image with landmarks, (b) triangulated reference face shape, (c)Spine image with landmarks, (d) triangulated reference spine shape.	177
Figure 7.3 Enhanced texture appearance modeling using ICA basis in Contourlet domain.....	181
Figure 7.4 Example of two-dimensional pose space using four proto-type poses.....	185
Figure 7.5 Block diagram of proposed spine x-ray image synthesis framework.....	186
Figure 7.6 Non-linear approximation of a spine x-ray image using Contourlets (top row), and Wavelets (bottom row). Reduction ratios: (a , e) 1: 200 (b , f) 1:125 (c , g) 1:100, (d , h) 1:40.	188
Figure 7.7 Non linear approximation of facial image by using Contourlets (Top row), and (Wavelets). Reduction ratios: (a , e) 1: 200 (b , f) 1:125 (c , g) 1:100, (d , h) 1:40.....	189
Figure 7.8 Examples of basis sets for different texture appearance models. (a) PCA-based Active Appearance Model, (b) Contourlet enhanced appearance model.....	190
Figure 7.9 Examples of localized texture variations obtained by using the new sparse texture model.	191
Figure 7.10 Reconstruction error of the new texture appearance model for different sizes of the training sets.	192
Figure 7.11 Examples of simulated x-ray spine images with various localized shape deformations being imposed (The texture and pose parameters are kept constant).....	193
Figure 7.12 Examples of simulated x-ray spine images with various texture variations being imposed (The pose and shape parameters are kept constant).....	194
Figure 7.13 Examples of simulated spine x-ray images and the corresponding segmentations. ..	195

List of Abbreviations and Symbols

X	Landmark-based shape vector
C	Multi-scale shape representation in wavelet packet domain
$I(x, y)$	A Grey-level image
φ	Basis function
φ_{best}	Best basis
c_i	Clinical variable of the i^{th} vertebra.
$T_{t_x, t_y, s, \theta}$	An operator to translate a shape by (t_x, t_y) , rotates it by θ , and scales it by S
Φ_{WICA}	Wavelet-ICA shape basis
Φ_{CTICA}	Contourlet-ICA texture appearance basis
$CT[.]$	Contourlet transform operator
rr	Repeatability Rate
ARR	Average Retrieval Rate
MSE	Mean Square error
TPR	True Positive Rate
FPR	False Positive Rate
NSCT	Non Sub-sampled Contourlet Transform
BB-ICA	Best basis ICA shape model
BB-PCA	Best basis PCA shape model
GT	Ground truth
FDP	Feature Discriminatory Power
FSL	Feature Spatial Locality
CTICA-AM	Contourlet-ICA based Appearance Model

1. Introduction

1.1 Motivation

Medical imaging is usually used to refer to a set of non-invasive techniques for probing the human body. The field of medical imaging dates back to the pioneer work of Professor Wilhem Ronteger and his seminal paper “On a new kind of Radiation.” Due to recent advances in imaging technologies, the scientific community has been challenged by the need to develop efficient, accurate and reproducible computer-aided methods to support medical experts in diagnostic and decision-making practices. The applications include early disease diagnosis, surgical planning, pathology-based image retrieval, and disease propagation and prediction.

A growing demand exists within the medical imaging community to develop computer-based methods that can achieve some sort of image interpretation and infer relevant quantitative measures. Image interpretation, always the ultimate goal of research in computer vision, includes the integration of a number of low-level and high-level tasks such as object detection, segmentation, recognition, registration and inferring.

In an attempt to achieve this long-standing goal, it seems logical to try to mimic the human visual system, which is the most amazing known image interpretation and analysis system. Researchers in cognitive studies believe that human vision is an active process. Our vision of the world is constructed from a combination of the retinal image and the prior knowledge we have about a scene. Traditionally, computer scientists have viewed the vision process as cascaded blocks of tasks such as registration, segmentation, and recognition. However, recent research efforts have indicated that this perception appears to be an over-simplification of the vision process. It is now believed that our visual system performs all these tasks in a simultaneous fashion that benefits from both low-level cues and high-level prior knowledge. Within the medical imaging community, this belief has increased the importance of prior knowledge-based models in various medical image analysis tasks.

This thesis is motivated mainly by two medical imaging tasks: the segmentation and analysis of non-rigid anatomical structures. We aim at developing model-based methods

for detecting deformations in medical images, and establishing relationship between these deformations and the various possible pathologies. Within a medical context, segmentation is the process of detecting anatomical structures in a complex image scene. Unlike man-made objects, anatomical structures are complex and subject to a large amount of variability. Statistical shape analysis is the process of locating changes between healthy and pathological structures. In this thesis, we are particularly interested in developing model-based methods to support practices related to the imaging of the human spine. Medical experts often examine hundreds of spine x-rays to look for various pathologies. Common pathologies include anterior osteophytes, disc space narrowing, and fractures. Detecting these pathologies normally requires the radiologist to perform the following tasks: locate the target structure in the image, accurately outline its edges, and employ prior knowledge to extract cues about the existence of a particular pathology. This process is both error-prone and time-consuming.

Thus, the work presented in this thesis is motivated by the need to develop new model-based methods to support various diagnostic practices conducted by medical experts. Inspired by studies about our human visual system, we aim to achieve the following goals:

- Developing novel statistical models that accurately capture prior knowledge about the shape and texture appearance of anatomical structures.
- Developing model-based methods for two spine-related tasks: the segmentation, and analysis of vertebral structures in medical images.
- Using the developed models of shape and texture to synthesize medical images with pathology-related deformations.

1.2 Medical Context

Throughout this work, we are interested in developing efficient statistical models of shape and appearance. We apply these models for the segmentation and analysis of vertebral structure in x-ray images to extract clinically relevant information. This section presents a brief overview of the medical context of the work presented in this thesis.

1.2.1 Anatomy of human spine

The human spine is made up of a number of vertebral bodies and joints referred to as inter-vertebral disks. Figure 1.1 shows the anatomy of the human spine. The spinal column is made up of a total of 33 vertebrae, separated by 23 inter-vertebral disks. It is divided into several regions: the cervical, thoracic, lumbar, sacrum, and coccyx spine. The vertebra is a box-like bony structure that protects the spinal cord. The inter-vertebral disk acts as a shock absorber between adjacent vertebrae. It helps minimize the movement between adjacent vertebral bodies, while allowing for more flexibility for the whole spine [1].

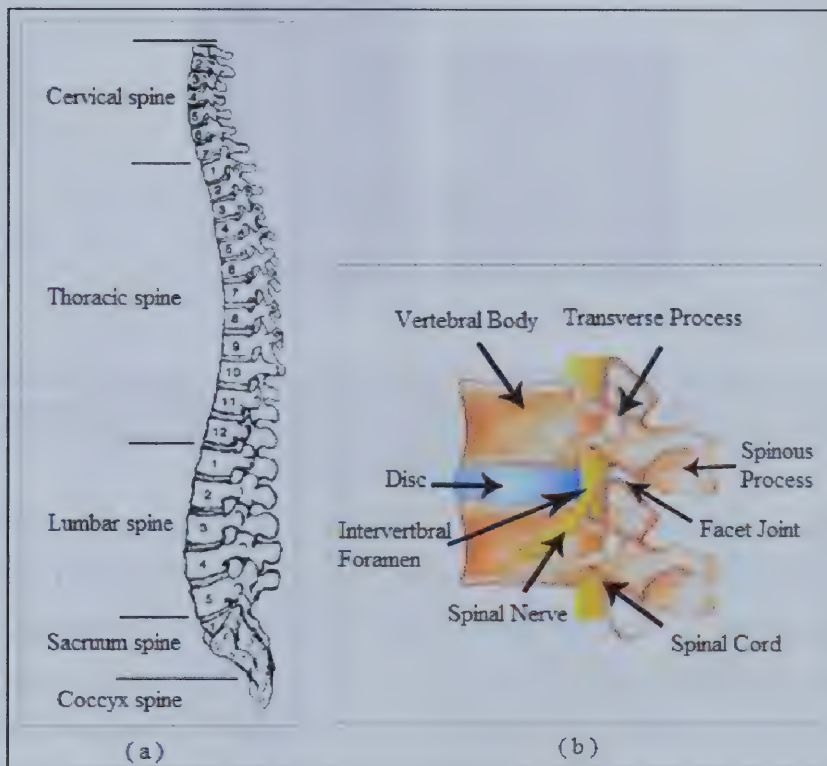


Figure 1.1 Anatomy of spine. (a) Lateral view of spine segments: Cervical, Thoracic, Lumbar, Sacrum, and Coccyx, (b) Lateral anatomy of a vertebra [1]

1.2.2 Common vertebral pathology

Medical experts often examine hundreds of spine x-ray images to look for various pathologies. Common pathologies of interests include osteophytes, disc space narrowing, and vertebral fractures. Figure 1.2 shows examples of x-ray images with different pathologies.

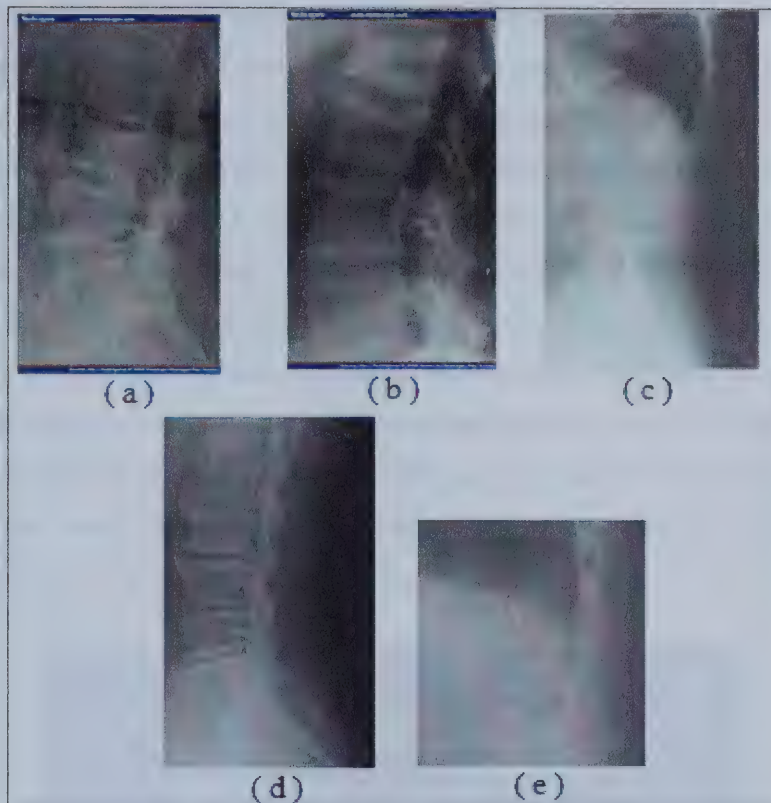


Figure 1.2 Lateral Radiographs for human spine with various pathologies (a) Lumbar spine with mild fracture in L3, (b) Lumbar spine with wedge fracture in L3, and crush in L2, (c) spine vertebrae with various fractures(e) Lumbar spine with osteophytes in L1,L2,L3, (f) Zoom-in showing disc space narrowing example [108]

Osteophyte pathology refers to bony growths on vertebral corners. Figure 1.2 (d) shows an example of the anterior osteophyte on lumbar vertebrae. Disc space narrowing, another common pathology in spine x-rays, refers to the change in inter-vertebral discs (degeneration). Numerous factors can contribute to this pathology, including age, gender, and injury. Disc space narrowing can be accompanied by back pain and nerve root compressions. Figure 1.2 (e) shows an example of vertebral structures with disc space narrowing pathology. Both osteophytes and disc space narrowing are often used as evidence of osteoarthritis disease. Osteoarthritis (degenerative joint disease) refers to the deterioration of joints in the body due to age, injury, or other diseases. Statistics show over 27 million patients have this disease in the United States [2].

Vertebral fracture is another important pathology that occurs due to a weakened vertebral body. Fractures are often detected by assessing the inclination of the vertebral endplates. Vertebral fracture is a continuous process requiring early diagnosis. A mild vertebral fracture is the most difficult to detect and appears as a small concave depression of the

endplate. A fracture can then progress to a wedge, which is a reduction in the anterior height of the vertebral body. A more severe fracture is the vertebral crush, which causes the vertebral body to start to lose its height. Figures 1.2(a) and 1.2(b) show examples of x-ray images with different types of fractures. Traditionally, fractures are assessed by using vertebral morphometry. A six-point representation of the vertebral body is used to conduct measures such as Anterior Height, Middle Height and Posterior Height [3]. In [4], Genant et al. proposed a semi-quantitative fracture grading system to assess the type and severity of vertebral fracture. Vertebral fractures are often used as evidence for Osteoporosis disease, characterized by low bone mass and structure deterioration of bone tissue, which result in fragile bones. According to [5], over 2 million of the vertebral fractures in the United States are due to Osteoporosis. It affects one in every two women and one in every four men over the age of fifty.

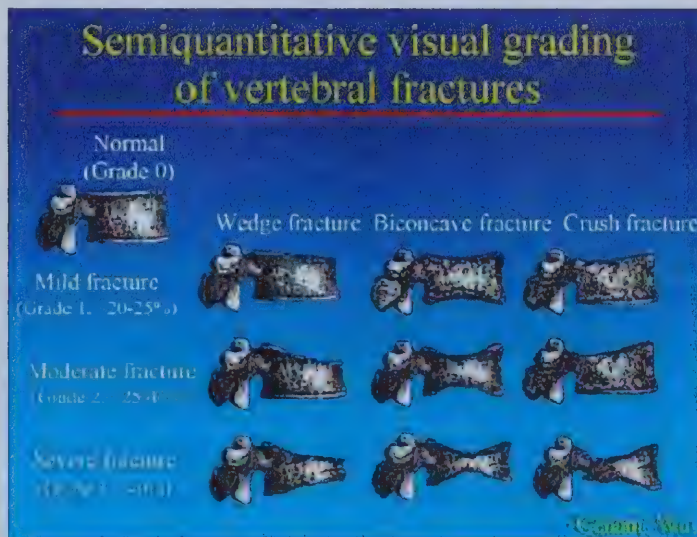


Figure 1.3 The Genant Semi-quantitative grading system [4]

Scoliosis is another disease often detected through assessment of vertebral morphometry and spine deformity. Scoliosis is a condition that causes lateral displacement of the vertebrae, resulting in curvature and rotation of the whole spine. It affects between 2-4 % of adolescents [6]. Scoliosis causes inclinations to the whole spine, as well as wedging to the individual vertebrae. Scoliosis is usually assessed by using the Cobb and wedge angles. The Cobb angle is obtained from the Posterior-Anterior Radiograph of the spine [7]. Vertebral and inter-vertebral wedge angles are two other famous indicators of Scoliosis [8]. The vertebral wedge angle is defined as the angular difference between the lower and upper endplates of the vertebral body. The inter-vertebral wedge angle is the

angular difference between the lower endplates of the top vertebra and the upper endplate of the bottom vertebra.

1.3 Challenges and Contributions

1.3.1 Challenges

This research deals with two challenging problems in medical imaging:

- i. Efficient statistical modeling of the shape and appearance variations of anatomical structures.
- ii. Application of these models in high-level spine-related tasks: the segmentation and analysis of vertebral structures.

Our work focuses on challenges such as those caused by the complexity of the anatomical structures, the wide range of variability that can take place, and the scarcity of ground truth data. The problems addressed in this thesis are outlined below.

Shape representation and modeling of anatomical structures

Statistical shape models try to faithfully represent the full range of biological variations in anatomical structures. However, the accurate representation and modeling of anatomical structures is a difficult task.

- Anatomical structures are complex and exhibit a wide range of local and global variations at various scales and frequencies. Thus, a rich representation of anatomical shape is crucial for the construction of efficient statistical shape models. This requirement is particularly true for articulated structures such as the human spine. Medical experts often examine hundreds of spine x-rays to determine the existence of various vertebral pathologies. Since most pathologies tend to be localized rather than global, the challenge is to find a suitable low-dimensional representation that can still capture subtle pathology-related information. This kind of representation would allow for the construction of better non-uniform shape models.

- An important application for shape priors is to constrain ill-posed problems such as segmentation and registration. Due to the limited availability of training samples and the complexity of anatomical shapes, statistical learning in the high-dimensional shape space constitutes an important challenge. Most shape priors suffer from either being too localized or too global. A localized shape prior uses only smoothness to constrain the solution. However, due to the uncertainty about the appearance of structures

in medical images, such priors might result in nonplausible solutions. On the other hand, an overly-constrained global prior fails to generalize well to cases unseen during the training phase. In this thesis, we focus on multi-scale methods to address this problem.

- We also focus on the clinical relevance of the model. The interpretability of model parameters has recently drawn much attention within the medical imaging community. This issue involves the ability of the model to extract modes of variations, which are physically intuitive and pathology related.

Statistical shape analysis in medical imaging

Statistical shape analysis refers to the ability to understand how anatomical structures tend to vary due to reasons such as the development of certain pathology, aging, or surgery. The medical imaging research community is constantly concerned with enhancing its ability to distinguish between the anatomy of healthy individuals and those with pathology. This enhancement would allow for the early detection and diagnosis of diseases. Moreover, it also opens the door for a wide range of applications such as surgical planning and pathology-related information retrieval. The efficient analysis and classification of anatomical structures continues to be a formidable task because such structures exhibit a wide range of inter-subject variability and structural differences among different pathologies. Traditionally, vertebral shape analysis has been performed by using a few points that outline vertebral bodies. Currently, shape-based methods are increasingly being used for better quantification of vertebral deformities. The challenge is to be able to develop robust shape analysis and classification methods with rich representations that can capture subtle pathology-related information.

Segmentation of vertebral structures in x-ray images

Segmentation of anatomical structures in medical images is a requirement for various applications such as computer-aided diagnosis, 3D visualization, and surgical planning. Computer-assisted segmentation methods aim at reducing the time and cost of the error-prone manual segmentation process. In this work, we are particularly interested in the segmentation of human vertebrae in digital radiographs. Accurate vertebrae segmentation plays a vital role in the proper assessment of various vertebral abnormalities. The goal is to develop efficient methods that can cope with various conditions such as pathology, weak edges, and noisy images. Developing such methods continues to be a difficult task

because of (i) poor image quality and noisy and incomplete data that affect the ability of low-level imaging procedures to accurately extract cues about the boundaries of the vertebral body; (ii) the complexity of the anatomical structure of the human spine and the wide range of possible variations, which tend to limit the ability to generalize to unseen examples; and (iii) the scarcity of training data, which also limits the performance of the segmentation method.

Modeling of texture appearance of anatomical structures

Appearance models attempt to capture the general texture properties of the object of interest. They are important for the interpretation of unseen images and the synthesis of new images. Traditionally, texture appearance models have been constructed by using a holistic Principle Component Analysis (PCA) basis extracted from a raw intensity pixel representation. Despite their potential benefits, traditional appearance models face a number of limitations. First, the scarcity of the training data tends to limit the model's accuracy due to the high dimensionality of the texture space. Second, modeling raw pixel intensities in texture appearance models becomes computationally infeasible in case of high-resolution images. This problem has led to a number of research efforts to develop enhanced appearance models. The challenges in texture appearance modeling include (i) the need to model the texture appearance in a domain that achieves high reduction rates while preserving higher-order details in the image such as boundaries and (ii) the fact that most appearance models use PCA for statistical modeling. Their use of PCA results in global texture modes of variations. However, most pathology is often related to local variations in both texture and shape.

1.3.2 Contributions

This thesis focuses on developing novel statistical models of shape and appearance to improve the support of various medical practices. We are concerned mainly with applying these models for the accurate characterization and segmentation of vertebral structures in x-ray images. This application is relevant for a wide range of clinical activities such as the diagnosis of various spine-related diseases (such as Osteoporosis, Osteoarthritis), surgical planning, and population studies of disease progression. Within this context, this thesis makes contributions in the following areas: (i) The construction of a descriptive multi-scale statistical shape model, (ii) efficient vertebral shape analysis and classification schemes, (iii) accurate prior-based segmentation of vertebral structures, and (iv) enhanced texture appearance modeling.

First, we attempt to tackle some of the challenges associated with modeling high-dimensional shape space by using a small training set. We aim at constructing an accurate shape model that captures both local and global pathology-related modes of variations by using a limited number of training samples. We propose a novel multi-scale statistical shape model that uses the concepts of sparsity, dimension reduction, and statistical independence to extract sets of multi-scale clinically relevant modes of deformations. Within a best basis selection framework, the proposed method benefits from the multi-scale and sparsifying nature of wavelet packets and the ability of Independent Component Analysis (ICA) to capture higher-order statistics. This method allows for the construction of a localized non-linear shape model with clinically-relevant modes of deformation. Our results have demonstrated that the new model outperforms the Active Shape Model and other related models in the literature, even in the presence of noise.

Second, we present a novel framework for the analysis and classification of vertebral structures by using pathology-related Wavelet-ICA shape parameters. For vertebral shape analysis, we use the parameters of the multi-scale shape model. Using statistical learning techniques, we are able to establish a statistically relevant relationship between the Wavelet-ICA modes of deformations and clinical information. Our results illustrate that significant differences were found between groups of healthy and pathological vertebrae. This finding allows for the anatomically meaningful interpretation of the multi-scale modes of variations. Furthermore, we propose a novel method for multi-vertebrae classification using Wavelet-ICA parameters and the Fisher classifier. We demonstrate the ability of the proposed framework to localize shape changes in vertebral structures related to a specific pathology. Thanks to the locality of the features, the new classification scheme can assess the condition of several vertebrae simultaneously. Our results show that the proposed framework achieves a better performance than that of analysis using the PCA-based representation of shapes.

Third, in an attempt to bring recent advances in computational harmonic analysis to the field of medical imaging, we present a novel vertebral segmentation algorithm using Contourlet-based local appearance profiles and a multi-scale shape prior. The new segmentation method aims to overcome challenging segmentation problems such as weak edges, noisy images, and the wide range of anatomical variations. At the core of the segmentation algorithm is a novel Contourlet-based salient point detection and matching

method. We employ the Non Subsampled Contourlet Transform (NSCT) to extract salient points of the vertebral body and build local appearance profiles for these points. Segmentation is then achieved via shape-driven matching of the salient points in the Contourlet domain. At a low level, the new salient point detector is used to detect evidence of the object of interest in the test image. At a high level, Contourlet-based local appearance profiles and the multi-scale shape prior are used to drive the segmentation process. Compared to Active Shape Models (ASM) [14] and related work in the literature, the proposed segmentation method achieved a high performance in terms of robustness to noisy images and generalization to unseen cases. We have also used the proposed Contourlet-based salient point detection algorithm in a Content-based medical image retrieval task. The proposed method benefits from the robustness and stability of the extracted salient points, and the information-rich Contourlet descriptors to retrieve images of similar classes. Promising results were obtained when applied to the IRMA (Image Retrieval in Medical Applications) database [125]. This work represents the first contribution from the Medical group at University of Alberta to the IRMA project.

Fourth, we tackle challenges related to the accurate modeling of texture and the scarcity of training data. We present a new sparse texture appearance model (CTICA-AM) based upon Contourlet Transform and Independent Component Analysis (ICA). The new model benefits from the non-linear approximation power of Contourlets to achieve high reduction rates. The statistical independence of Independent Component Analysis (ICA) is employed to capture localized texture variations. The proposed Contourlet-enhanced Texture Appearance Model is shown to outperform related appearance models in the literature, especially at high reduction rates. We also describe a general framework that integrates developed shape and texture models in a unified framework for the synthesis of photo-realistic pathological x-ray images.

1.4 Thesis Organization

The rest of the thesis is organized as follows:

Chapter 2 presents some of the background material essential for understanding the methods developed in this thesis. The chapter also provides a brief review of the most relevant work in the literature.

Chapter 3 presents an overview of the recent advances in multi-scale image analysis tools. The chapter also discusses the suitability of various multi-scale tools for medical image analysis and interpretation.

Chapter 4 presents the details for a novel multi-scale statistical shape model. Experimental results are also provided for the quantitative and qualitative evaluation of the new shape model.

Chapter 5 presents the details for a new vertebral shape analysis and classification framework. Experimental results are provided for the evaluation of the proposed method.

Chapter 6 presents details for a new salient point detection and matching algorithm. The chapter also presents novel vertebrae segmentation and retrieval methods.

Chapter 7 presents a new Contourlet-enhanced appearance model. The chapter also presents a general framework for the synthesis of realistic spine x-rays with various pathological deformations.

Chapter 8 summarizes the work presented in this thesis. The chapter also discusses some of the remaining challenges and recommendations for future work.

2. Background and Related Work

2.1 Introduction

Decades of research in computer vision and cognitive studies have led to a gradual shift from “bottom-up” to “top-down or “model-based” strategies for image analysis and interpretation. In “bottom-up” strategies, images are analyzed by using pure low-level procedures by looking for local structures such as edges, contours, and regions. The evidence is then grouped together and used to identify objects of interest. While this approach might be able to achieve limited success in controlled environments, it has proven to be difficult to use and prone to failure in more challenging situations such as medical images. On the other hand, “top-down” strategies tend to use prior-knowledge about what is expected in the image. Models of prior-knowledge are matched to the data in the test image. Model-based vision systems often employ a combination of bottom-up and top-down approaches. Over the past decades, model-based vision methods have been successively applied to various image analysis and computer vision tasks [9]. Specificity and generalizability are the two major challenges in any model-based vision system, particularly in medical image analysis, because of the noise in the images and the complexity of the anatomical structures of interest.

The segmentation of anatomical structures is a fundamental step in almost any medical image analysis task. Segmentation is the process of outlining the anatomical structure of interest for further processing. Unlike man-made objects, anatomical structures are complex and subject to a wide range of variability. In an attempt to address these challenges, deformable models have been introduced and extensively used for the segmentation and analysis of anatomical structures. Deformable models attempt to fit a new image by applying constraints on the topology of the structure’s boundary. The first deformable model was introduced in 1987 by Kass et al. [28]. Over the past two decades, many extensions and modifications have been proposed in the literature.

This thesis is motivated mainly by the desire to perform two imaging tasks: the segmentation and quantification of anatomical structure. We are also interested in establishing the relationship between deformations and various pathologies. The rest of this chapter presents some of the background material essential for understanding the methods developed in this thesis. We start with an overview of the basics of shape

representation in medical images. In section 2.3, we present a brief review of various deformable models. In sections 2.4 and 2.5, we present the basics of the Active Shape Model (ASM) and Active Appearance Model. Section 2.6 discusses some of the most relevant work in the literature.

2.2 Shape representation

Shape representation plays a critical role in the success of higher-level imaging tasks. In [10], Kendall provides a definition of shapes as follows:

Shape: Two geometric objects have the same shape, if one can be mapped onto the other one by a similarity transformation. Similarity transformation includes translation, rotation, and scaling.



Figure 2.1 Example of 2D shape under different similarity transformations.

The medical imaging literature contains a wealth of methods for efficient shape representation that support a variety of higher level tasks such as segmentation, analysis and modeling. Generally, shape representation methods can be classified into two categories: parametric and non-parametric. Examples of parametric shape representations are spline-based representations [11] and Fourier descriptors [12]. Examples of non-parametric shape representations are landmark-based methods [13, 14], medial axis representation [15], and implicit level set representation [16].

2.2.1 Parametric representation

This section presents the basics of parametric shape representation. In chapter 4, we present a novel multi-scale parametric shape representation by using a best-basis selection framework. A parametric shape describes the shape with a set of parameters. In general, a parametric shape descriptor represents the boundary of a contour by using a set

of ordered control points. The control points are parameterized by a parameter (s) that varies from 0 to 1:

$$\vec{C}(s) : [0,1] \rightarrow R^2 \quad (2.1)$$

$$\vec{C}(s) = [x(s), y(s)] . \quad (2.2)$$

In its simplest form, a closed curve of control points is represented by using a polygonal approximation with the control points at the vertices of the polygon and connected by straight lines.

2.2.1.1 Spline-based representation

Spline-based representation offers a smoother representation of the contour in comparison to the polygonal approximation, as shown in Figure 2.2. In B-spline shape representation, a piece-wise polygonal curve is represented as a linear combination of control spline basis functions:

$$\vec{C}(s) = \sum_{n=0}^{N-1} \alpha_n \varphi_n(s) , \quad (2.3)$$

where :

α_n : Coefficients of the representation

φ_n : Bases functions of the spline

$\vec{C}(s)$ is a parameterized curve with parameter s .

Spline shape representation decomposes the shape into a combination of spatially localized polynomial basis functions.

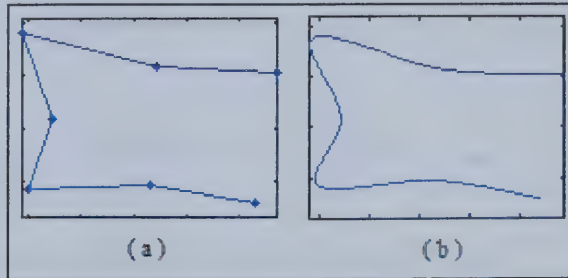


Figure 2.2 Parametric shape representation using (a) polygonal approximation, (b) B-Spline approximation

2.2.1.2 Fourier Descriptor

Another famous parametric shape representation is the Fourier descriptor [12], which represents a contour by using a linear combination of Fourier bases.

Let $\vec{C}(s) = [0,1] \rightarrow R^2$ be a 2D continuous curve. The Fourier descriptor is then calculated as follows:

$$\vec{C}(s) = \sum_n \alpha_n e^{i(2\pi/L)ns} \quad (2.4)$$

$$\alpha_n = \frac{1}{L} \int_0^L \vec{C}(s) e^{-i(2\pi/L)ns} ds, \quad (2.5)$$

where α_n is the Fourier coefficient.

In medical imaging, Fourier descriptors have been used to represent anatomical structures. In [12], Staib et al. used Fourier descriptors to construct a shape prior for image segmentation. In [17], the authors described a Fourier-based shape prior to constrain deformations in deformable Active Contours known as Snakes. In [18], Ghassan et al. described a statistical shape prior based upon Discrete Cosine Transform.

2.2.2 Non-Parametric representation

2.2.2.1 Landmark-based representation

A landmark-based representation describes the contour of the object of interest by using a finite set of points on the boundary which are referred to as landmarks. Mathematically, an n -point shape in d -dimensional space is represented by concatenating each dimension into a dn -dimension vector. Thus, a $2D$ shape is represented as

$$\mathbf{x} = [x_1 \ x_2 \dots x_n \ y_1 \ y_2 \dots y_n]^T \quad (2.6)$$

Where x_i and y_i are the x- and y- coordinates of the i^{th} point.

Figure 2.3 shows the outline of several brain structures by using landmark-based representation. The landmarks are most commonly placed manually by experts who can

utilize their knowledge of anatomy and potential pathologies. Generally, landmarks can be either anatomical, mathematical, or pseudo landmarks:

- (i) Anatomical landmarks: Correspond to biologically relevant points.
- (ii) Mathematical landmarks: Correspond to mathematical or geometric property of boundary.
- (iii) Pseudo landmarks: Used to obtain dense representation of shape by placing landmarks at equally spaced distances between other landmarks.

Landmark-based shape representation is used in the Active Shape Models (ASM) proposed by Cootes et al. [12]. Statistical shape modeling using landmark-based representation requires establishing the correspondence among the landmarks in the whole training set. Within this context, correspondence means that each landmark is placed at the same anatomical location over the whole training set. Several studies have suggested new methods to automate the process of identifying the landmarks and establishing the correspondence among them [19, 20].

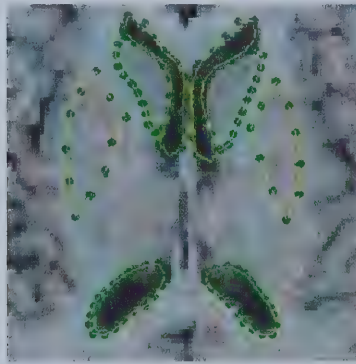


Figure 2.3 Outline of brain structure by using landmark-based representation (from [21])

2.2.2.2 Medial axis-based representation

According to [15], the medial axis of a shape is defined as a set of the singularity points of the distance transform that lies inside the shape of interest. Medial-axis representation describes the skeleton of the object of interest. Recent physiological studies have indicated that medial profiles play an important role in the Human Visual System (HVS). In the field of medical imaging, medial axis representation has been successfully used for segmentation and analysis tasks [22, 33]. Figure 2.4 shows a typical simulated corpus callosum and its corresponding medial axis representation.



Figure 2.4 (a) Simulated corpus callosum image, (b) medial-axis representation (From [22])

2.2.2.3 Implicit Level-Set representation

Level-set is a method for shape representation that implicitly accounts for a change in topology. A level-set representation describes the shape as the zero-level of a higher dimensional function. Mathematically, a level-set is described as follows:

A 2D curve $\vec{C}(s) = [0,1] \rightarrow R^2$ is represented as being embedded as zero-level of a specific function in R^3 whose height is sampled at regular intervals on the xy grid [16]. Level sets are typically used in medical image segmentation [24]. In [25], Leventon et al. suggested adding shape priors to level sets for better segmentation accuracy.

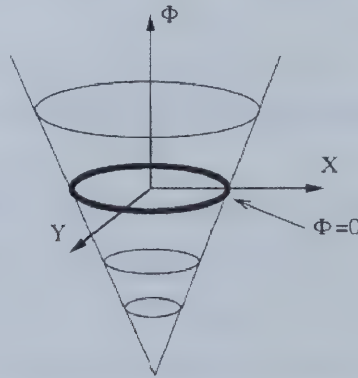


Figure 2.5 Implicit Level set representation. A curve is given by a set of points onto the level set function Φ

2.3 Deformable Models

Formally, a deformable model can be defined as follows [26]:

Deformable model: A model which under an implicit or explicit optimization criterion deforms a shape to match a known object in a given image.

In general, deformable models use prior knowledge about the object to facilitate the extraction of the object boundary. In [27], Jain et al. classified deformable models as being either free-form or parametric. Free-form models employ local shape constraints (e.g., smoothness) as the only prior knowledge. On the other hand, parametric models use global shape constraints. Figure 2.6 summarizes Jain's classification of the deformable models.

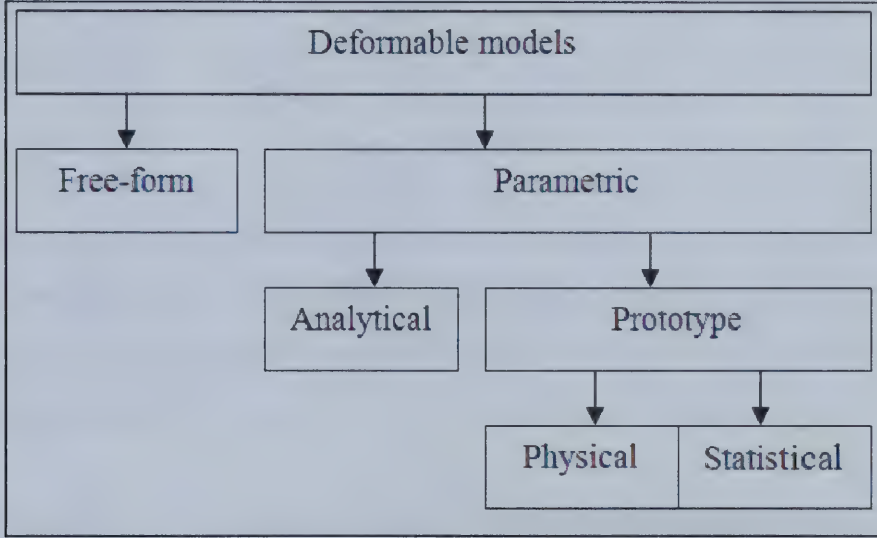


Figure 2.6 Classification of deformable models

2.3.1 Active Contours (Snakes)

In [28], Kass et al. introduced the most popular deformable model known as Active Contours (or Snakes). Snakes are free-form deformable curves that move freely under the influence of a combination of forces: (i) an internal force that imposes local constraints such as smoothness and (ii) an external force that encourages movements towards desired image features such as edges.

Mathematically, Snakes are formulated as an energy-minimization problem [29]. A Snake is a parameterized contour embedded in the image plane $(x, y) \in R^2$. Let $v(s) = [x(s), y(s)]^T$ represents the outline of the contour, where $x(s)$, and $y(s)$ are the coordinate functions, and $s \in [0, 1]$ is the parametric domain. The shape of the contour is defined by the energy function $E(v)$:

$$E(v) = S(v) + P(v), \quad (2.7)$$

where $S(v)$ represents the energy of the contour, and $P(v)$ represents the energy that couples the Snake to the image. The deformation of the contour is characterized by the following equation:

$$S(v) = \int_0^1 w_1(s) \left| \frac{\partial v}{\partial s} \right|^2 + w_2(s) \left| \frac{\partial^2 v}{\partial s^2} \right|^2 ds, \quad (2.8)$$

where $S(v)$ characterizes the deformation of the contour, $w_1(s)$ controls the tension of the contour, and $w_2(s)$ determines its rigidity. The potential energy is given by the following equation:

$$P(v) = \int_0^1 \Psi(v(s)) ds. \quad (2.9)$$

$P(v)$ is a potential energy function that couples the contour to the image plane. $P(v)$ is usually designed to have local minima coincide with the intensity extrema in the image plane. A Snake with a potential function $\Psi(x, y) = -c |\nabla[G_\sigma * I(x, y)]|$ will be attracted to the intensity edges in the image.

The final shape of the contour corresponds to the minimum of equation 2.7. Using Calculus of variations, the contour, $v(s)$, that minimizes equation 2.7 must satisfy the Euler-Lagrange equation:

$$-\frac{\partial}{\partial s} \left(w_1 \frac{\partial v}{\partial s} \right) + \frac{\partial^2}{\partial s^2} \left(w_2 \frac{\partial^2 v}{\partial s^2} \right) + \nabla P(v(s, t)) = 0. \quad (2.10)$$

Equation 2.10 represents the balance between the internal and external forces when the contour rests at equilibrium. Figure 2.7 shows an example of applying Snakes to a segmentation problem. Generally, free-form models have no global shape constraints. Thus, they are best suited to aid in manual segmentation tasks. It is now accepted that free-form models suffer from a number of limitations. First, they need to be initialized close to the structure of interest. Second, they use only the smoothness constraints of the contour, and this limitation can result in non-plausible shapes. Over the past two decades, many improvements have been suggested to the original Snake model. In [30], Cohen suggested using an internal inflation force in the Snake model to increase its sensitivity to spurious edges. In [31], Xu et al. proposed a Gradient vector field-based snake. A multi-scale formulation of the snake model was introduced in [32]. In this formulation, the

Snake is allowed to come to equilibrium on a very low-scale energy function and then slowly increases the scale. In an attempt to deal with changes in topology, Malladi et al. proposed a deformable model using level sets [11]. In [18], Hamarneh et al. proposed to embed both local and global shape constraints within the free-form model.

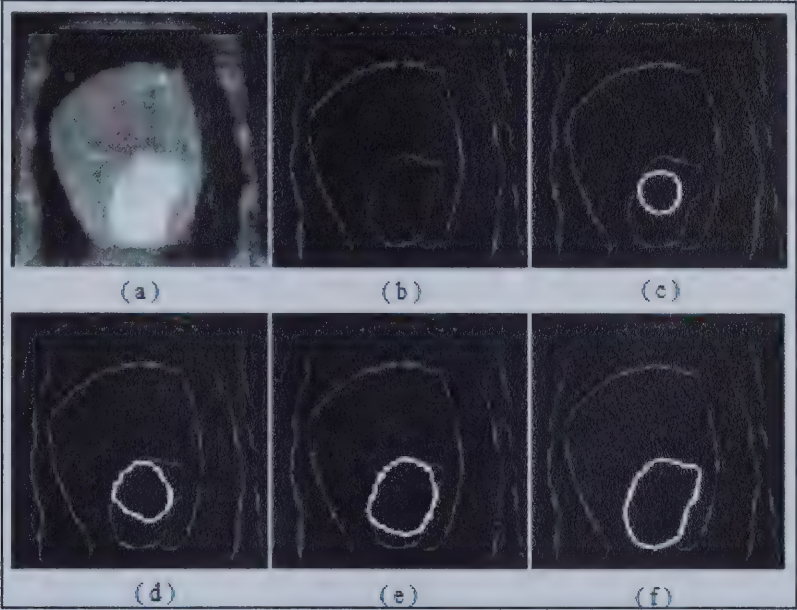


Figure 2.7 Example of Snake-based segmentation (a) Intensity CT image of heart left ventricle (b) edge detected image (c) initial contour (d-f)deformable contour moving toward left ventricle boundary (from [29])

2.3.2 Physical models

Physical deformable models are a class of deformable models in which a single example (prototype) is used to build the parametric model. The prototype is allowed to vary according to some physical model. Park et al. [33] and Pentland et al. [34] presented a model using the Finite Element method. Variations are defined as the vibration modes of the model. Figure 2.8 shows examples of the Finite Element modes of vibrations for a square and a lumbar spine. A common criticism of physical models in medical imaging is that nothing guarantees that the generated modes of variations will correspond to plausible deformations.

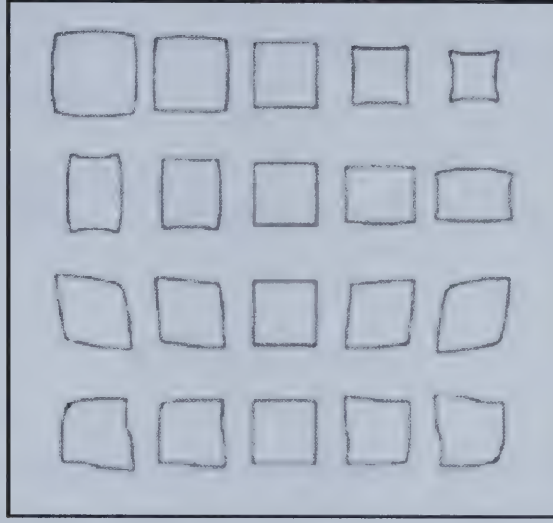


Figure 2.8 Modes of vibrations of Finite Element model for a square shape

2.3.3 Statistical shape models

Statistical shape models aim at capturing acceptable patterns of variations in a class of objects by using a training set of examples. In [12], Staib and Duncan suggested using elliptic Fourier descriptors to represent the contour of the object. The Fourier coefficients are trained by using a Gaussian distribution as the prior probability. The optimal estimate of the boundary is then derived by using the Bayesian decision rule. This model represents one of the earliest statistical shape models in medical imaging. However, using the Fourier basis tends to limit the model's ability to capture localized variations. Cootes et al. [35] presented the seminal work involving the Active Shape Models (ASMs). The Active Shape Model is a statistically-based parametric deformable model that is used to outline objects of interest in the image. The model is built up from a representative average shape and a weighted sum of modes of variations learnt during the training phase. Segmentation is achieved by allowing the model to deform only in a way implied by the training set.

In the subsequent sections, we present the details of the Active Shape Models and the Active Appearance Models.

2.4 Active Shape Models

The Active Shape Model (ASM) was introduced by Cootes et al. in [35]. A detailed description of the model and its applications was later presented in [36]. The Active Shape Model is a deformable model that globally constrains deformations to valid instances of the object of interest. This global shape constraint tends to improve the

specify of the model and to provide better segmentation. The ASM algorithm has three main parts: (i) The point distribution model (PDM), (ii) The local appearance model, and (iii) the model search procedure.

2.4.1 Point Distribution Model

The contour of the object is represented by using a fixed number of labelled landmarks p_i :

$$p_i = (x_i, y_i), i = 1 \dots N. \quad (2.11)$$

Prior to modeling the shape variations in the training set, they are aligned with respect to translations, rotations, and scaling by using the Procrustes analysis method. The following summarizes the alignment process [38]:

- i. Choose an example shape as the mean shape.
- ii. Align all shapes to the mean.
- iii. Re-calculate the estimate of the mean by using the aligned shapes.
- iv. In case the new mean estimate differs significantly from the other one, repeat from step (ii).

Non-rigid variations are then modelled by using the Point Distribution Model (PDM). The PDM uses Principle Component Analysis (PCA) to capture the second-order statistics in the training set. The Point Distribution Model represents a shape as a mean shape plus a weighted sum of basis functions defined as the eigenvectors of the covariance matrix. Mathematically, the PDM is defined as follows.

The mean shape is

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (2.12)$$

The Covariance Matrix is

$$S = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})^T. \quad (2.13)$$

The eigenvectors φ_k and eigenvalues λ_k of S are then computed:

$$S\varphi_k = \lambda_k \varphi_k. \quad (2.14)$$

The matrix of the eigenvectors $\Phi = [\varphi_1 \varphi_2 \dots \varphi_N]$ represents the modes of variations of the shapes. Usually, variations in the training set can be explained by using a small number of modes (i.e., eigenvectors). Thus, we can obtain a good estimate of the shape by retaining the eigenvectors φ_i that correspond to the t largest eigenvalues:

$$x = \overline{x} + \Phi \mathbf{b} \quad (2.15)$$

$$\Phi = [\varphi_1 \varphi_2 \dots \varphi_t], \quad (2.16)$$

where $\mathbf{b} = [\alpha_1 \alpha_2 \dots \alpha_t]^T$ is the vector of model parameters. The variance of α_k over the training set can be shown to be equal to the eigenvalue λ_k . Thus, we can obtain plausible shapes by restricting the model parameters to vary within a range of $\pm m\sqrt{\lambda_i}$. Figure 2.9 shows two modes of variations for a Point Distribution Model of hand shapes.



Figure 2.9 First two modes of variations of PDM model for hand shapes

2.4.2 Local appearance model

During the training phase, a local appearance profile is constructed for every landmark point of the contour of interest. Given a training set of N_p images, the local appearance profile is constructed as follows:

For every image in the training set (I_j), we extract a gray level profile $\mathbf{g}_{ij}$ for every landmark point P_i of the contour of interest. The grey level profile is a vector of

intensity values sampled from the linear profile passing by the landmark and perpendicular to the shape:

$$\mathbf{g}_{ij} = [g_{ij0}, g_{ij1}, \dots, g_{ijN-1}]^T, \quad (2.17)$$

where $\mathbf{g}_{ij}$ is a profile of the i^{th} landmark point extracted from the image I_j using N samples.

In order to obtain invariance to uniform scaling, we obtain the normalized derivative of the grey level profile:

$$\mathbf{y}_{ij} = \frac{\mathbf{dg}_{ij}}{\sum_{k=0}^{N-2} |\mathbf{dg}_{ij}|} \quad (2.18)$$

$$\mathbf{dg}_{ij} = [g_{ij1} - g_{ij0}, g_{ij2} - g_{ij1}, \dots, g_{ijN-1} - g_{ijN-2}]^T, \quad (2.19)$$

where $\mathbf{y}_{ij}$ is the normalized derivative of the grey level profile, and $\mathbf{dg}_{ij}$ is the derivative of the grey level profile.

Given a set of N_p normalized derivatives of the grey level profiles, the local appearance profile for the i^{th} landmark point is then defined as the mean profile $\overline{\mathbf{y}}_i$ and the covariance matrix $\mathbf{S}_i$:

$$\overline{\mathbf{y}}_i = \frac{1}{N_p} \sum_{j=1}^{N_p} \mathbf{y}_{ij} \quad (2.20)$$

$$\mathbf{S}_i = \frac{1}{N_p} \sum_{j=1}^{N_p} (\mathbf{y}_{ij} - \overline{\mathbf{y}}_i)(\mathbf{y}_{ij} - \overline{\mathbf{y}}_i)^T. \quad (2.21)$$

2.4.3 Model search procedure

Segmentation is then achieved by fitting the learnt model to a new image by estimating the model parameters ($\mathbf{b}$) and the pose parameters (t_x, t_y, s, θ) that best fit the image:

$$\mathbf{x} = T_{t_x, t_y, s, \theta}(\overline{\mathbf{x}} + \phi \mathbf{b}). \quad (2.22)$$

$T_{t_x, t_y, s, \theta}$ is a function that translates the shape by (t_x, t_y) , rotates it by θ , and scales it by s . We start with an approximate pose estimate $(t_{x_0}, t_{y_0}, s_0, \theta_0)$, and the model mean

shape ($\bar{\mathbf{x}}$). The pose and model parameters are then estimated iteratively so that the model best fits the new image. The search iteration consists of the following two tasks:

(i) For each landmark point, examine a search profile along a line perpendicular to the contour and passing by the landmark point. Find the location that best fits the learnt local appearance profile of this landmark. This is the location that minimizes the Mahalanobis distance ($h(\mathbf{y}_s)$). The output is a new shape estimate based upon the local appearance profiles:

$$h(\mathbf{y}_s) = (\mathbf{y}_s - \bar{\mathbf{y}}_k)^T \mathbf{S}_{y_k}^{-1} (\mathbf{y}_s - \bar{\mathbf{y}}_k) , \quad (2.23)$$

where $\mathbf{y}_s$ is the local appearance profile for the current position in the search profile, and $\bar{\mathbf{y}}_k, \mathbf{S}_{y_k}$ is the local appearance model (mean and covariance matrix) for the k^{th} landmark.

(ii) Update the pose and model parameters ($t_x, t_y, s, \theta, \mathbf{b}$) to best match the new shape estimate from (i):

$$\mathbf{b} = \boldsymbol{\phi}^T (\mathbf{x}_s - \bar{\mathbf{x}}) \quad (2.24)$$

$$\tilde{\mathbf{x}} = T_{t_x, t_y, \theta, s} (\bar{\mathbf{x}} + \boldsymbol{\phi} \mathbf{b}) . \quad (2.25)$$

More details for updating the parameters of the ASM model are given in Appendix B.

2.5 Active Appearance Model

The Active Appearance Model (AAM) is a generative model that can synthesize images similar to those in the training set [37]. Typically, Active Appearance Models are used for image interpretation and synthesis applications. Using a training set of images, we model both the shape and texture (pixel intensity) of the object of interest. Image interpretation tasks can then be achieved by adjusting the model parameters to fit the new unseen image.

2.5.1 Shape model

Shape variations across the training set are estimated by using the Point Distribution Model (PDM) described in section 2.4. The outcome is an estimate of a mean shape and a set of PCA-based basis functions that define the modes of shape variations:

$$\mathbf{x} = \overline{\mathbf{x}} + \boldsymbol{\varphi}_s \mathbf{b}_s, \quad (2.26)$$

where $\overline{\mathbf{x}}$: Mean shape,

$\boldsymbol{\varphi}_s$: Matrix of shape modes of variations,

$\mathbf{b}_s$: Vector of shape parameters.

2.5.2 Texture model

The Active Appearance Model (AAM) also models the texture appearance of the object.

First, we sample the pixel intensities within the convex hull of the shape of the object. All

samples are then warped to the estimated mean shape ($\overline{\mathbf{x}}$) by using the Delaunay triangulation method [39]. Next, Principle Component Analysis (PCA) is used to model the texture appearance, as well as the combined variations between the shape and texture:

$$\mathbf{g} = \overline{\mathbf{g}} + \boldsymbol{\varphi}_g \mathbf{b}_g, \quad (2.27)$$

where $\overline{\mathbf{g}}$: mean normalized grey-level vector,

$\boldsymbol{\varphi}_g$: Modes of intensity variations,

$\mathbf{b}_g$: grey-level model parameters.

In order to account for any potential correlation between the shape and the grey-level model parameters, Principle Component Analysis (PCA) is applied to a concatenated vector of $\mathbf{b}_s$ and $\mathbf{b}_g$:

$$\begin{pmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix} = \mathbf{b} = \begin{pmatrix} \mathbf{Q}_s \\ \mathbf{Q}_g \end{pmatrix} \mathbf{c} = \mathbf{Q} \mathbf{c}, \quad (2.28)$$

where $\mathbf{Q}_g$: Basis matrix of texture model,

$\mathbf{Q}_s$: Basis matrix of shape model,

$\mathbf{Q}$: Basis matrix of combined appearance model,

$\mathbf{c}$: Parameter vector of combined appearance model.

The complete model can then be formulated as follows:

$$\mathbf{x} = \overline{\mathbf{x}} + \boldsymbol{\varphi}_s \mathbf{W}_s \mathbf{Q}_s \mathbf{c}, \quad \mathbf{g} = \overline{\mathbf{g}} + \boldsymbol{\varphi}_g \mathbf{Q}_g \mathbf{c}. \tag{2.29}$$

A new image can be synthesized as follows:

- Generate a shape-free grey level image by using the appearance parameter vector.
- Warp the new shape-free image to match the landmarks of the target shape.

Figure 2.10 shows the first two modes of variations for an active appearance model of the lumbar spine.

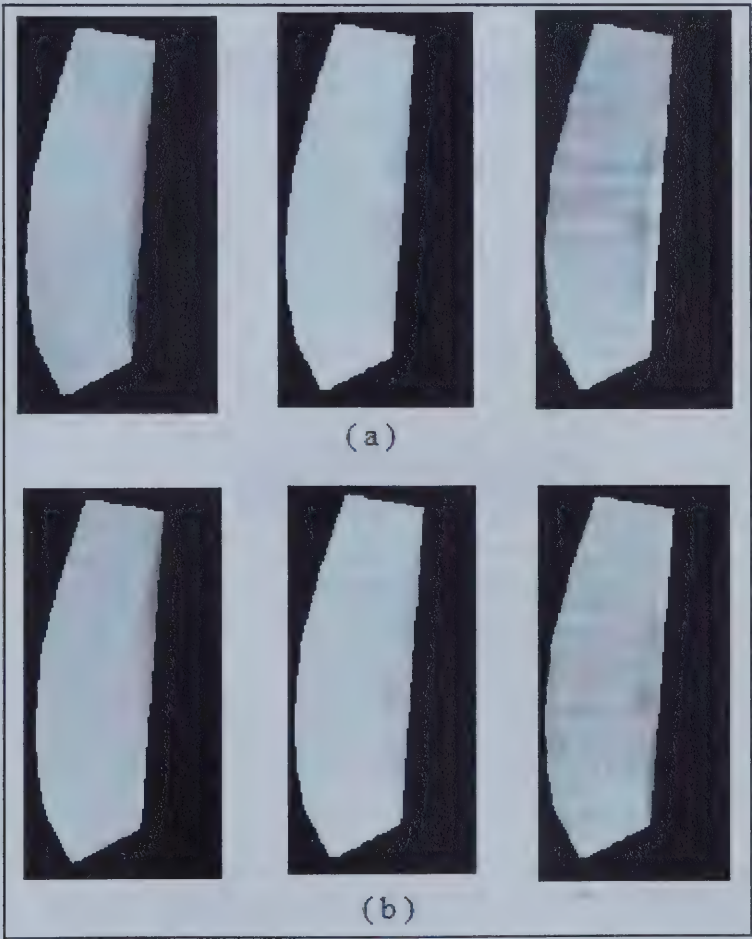


Figure 2.10 First two modes of appearance model of the lumbar spine

2.5.3 Image Search

The Active Appearance model is then matched to unseen images by using least square criteria. The algorithm minimizes the difference between the sample from the searched

image ($\mathbf{g}_s$) and the synthesized sample image ($\mathbf{g}_m$) by varying the model parameters which control the shape, texture, and pose. For more details, refer to [38].

2.6 Related Work

This thesis describes new methods for the statistical modeling of the shape and appearance of anatomical structures. We apply these models for detecting and quantifying of vertebral structures in x-ray images. Figure 2.11 shows a block diagram of the general framework of the work presented in this thesis. This section presents a brief review of the most relevant work from the literature.

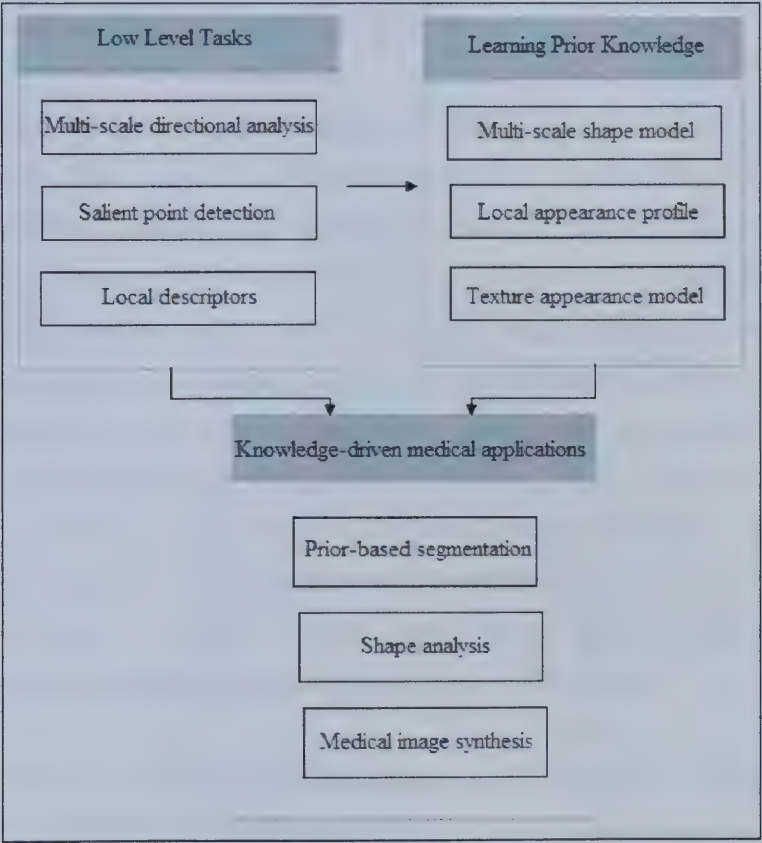


Figure 2.11 General framework for the work presented in the thesis

2.6.1 Statistical Shape Modeling

Statistical shape models try to faithfully represent the full range of biological variations in anatomical structures. These models are now widely used as prior knowledge in a wide range of applications such as segmentation, registration, and surgical planning. However, the accurate modeling of anatomical structures is a challenging task due to shape complexity and the wide range of possible local and global variations.

Principal Component Analysis (PCA) has been widely used in the literature for modeling shape variations. Nevertheless, many researchers now understand that PCA basis functions are global in nature, and fail to generalize well to complex unseen deformations. To overcome the limitations of PCA-based shape models, several approaches have been suggested in the literature [40].

Pentland and Sclaroff, [33], suggested a shape model based upon the mechanical modes of deformations. This model eliminates the need for a large training set. Cootes and Taylor [41] proposed a hybrid shape model that combines both statistical and mechanical modes to produce better generalization. The main problem with physical models is that the resulting deformation might not reflect realistic shape variations. Wang et al. [42] suggested using a mixed covariance matrix, which is defined as a weighted sum of the estimated covariance matrix plus another matrix that imposes smoothness. Cootes and Taylor [43] suggested modeling the shapes in the training set as a mixture of Gaussians. The authors showed that their model could better capture non-linear variations than traditional Active Shape Model (ASM). In [44], Bowden et al. proposed a new method to overcome the linearity in the PCA model. Data are first clustered into a number of clusters. A separate PCA model is then constructed for each cluster.

The work of Davatzikos et al. [45, 46], Mahmoud [40], and Nain et al. [47, 48], is more related to the methods presented in this thesis. In [45], Davatzikos et al. proposed a hierarchical shape prior using the wavelet transform and PCA. The authors suggested by using the Discrete Wavelet Transform (DWT) to construct a hierarchical representation. Principal Component Analysis (PCA) is then used to capture the statistics of the coefficients. In [40], Mahmoud suggested by using a wavelet packet basis set. The coefficients are then modeled using Principal Component Analysis (PCA). In [47, 48],

Nain et al. proposed an approach to model caudate nucleus shapes by using spherical wavelets and Principal Component Analysis (PCA).

In an attempt to overcome the inherent limitations in PCA-based models, a recent trend has involved the search for a better deformation basis set more suited for modeling anatomical structures. Instead of finding the basis that maximizes the variance in the training set, some researchers proposed using a basis set that maximizes the statistical independence in the training set. The Independent Component Analysis (ICA) method was first proposed for shape modeling applications in [49]. An ICA-based shape model represents a shape as a mean shape plus the weighted sum of the independent modes of variations. Rogez et al. [50] used an ICA-based shape model for the segmentation of a human figure. Uzumcu et al. [51] suggested using Independent Component Analysis (ICA) for modeling shapes in medical applications. Koikkalainen et al. [52] proposed a combined ICA- and PCA- based model for the segmentation of magnetic resonance images. To achieve more interpretable localized models, Sjostrand et al. proposed the use of Sparse Principal Component Analysis (SPCA), which imposes a sparsity constraint onto the PCA basis [53, 54]. Chennubhotla et al. proposed a method for extracting multi-scale structure from data by using sparsePCA [55]. In [111], Olafsdottir et al. compared the performance of PCA, ICA, and SPCA in the statistical deformation modeling of a Crouzon mouse. In [122], Suinesiaputra et al. presented a method for detecting abnormal cardiac contraction in short-axis MR images by using ICA.

2.6.2 Statistical Appearance Modeling

Generative models are often used in medical imaging to constrain solutions for various complex ill-posed problems. Statistical models of appearance are popular examples of generative models. In chapter seven, we present a novel method for enhanced modeling of texture appearance, with an application to medical image synthesis. This section summarizes some of the related work in the literature.

In [37], Cootes et al. presented the Active Appearance Model (AAM) as a new method to capture combined variations in shape and texture. A review of the basics of the Active Appearance Model was presented in section 2.5. Since the introduction of Active Appearance Models, the modeling of texture appearance in medical images has been performed mainly by using a PCA basis. However, using the PCA basis to model texture

appearance has a number of limitations. The PCA basis captures only the global modes of texture variation, while recent studies have emphasized the role of locality in medical imaging. Most pathologies are often related to local deformations in both geometry and texture [55]. In order to overcome the limitations of Active Appearance Models, a number of researchers have tried to achieve locality in appearance modeling. Roberts and Cootes [56] proposed a method to capture localized texture variations by using linked active appearance models. These models are used for the segmentation of lumbar vertebrae in digital radiographs. Delac et al. [57] proposed a statistical appearance model using Independent Component Analysis (ICA). This model is used for face recognition. In [59], Zhang et al. presented a method for modeling human faces with parameterized local shape morphing. In [60], Stegmann et al. presented a method for the sparse modeling of landmark and texture variability by using the Orthomax criteria. The model generates localized shape and texture variations.

Another challenge that faces traditional intensity-based appearance models is the high resolution of medical images. The higher the image resolution, the more unfeasible it becomes to model texture appearance by using raw intensity pixels. Researchers have suggested various methods for capturing texture variations in other compressed spaces. The challenge is to achieve high reduction rates without degrading the performance of the model. In [61], Wolstenholme and Taylor proposed using Haar-based wavelet compression in Active Appearance Models. A compression ratio of 20:1 was accomplished while maintaining acceptable accuracy in the model. In [62], Stegman and Cootes extended the work in [61] by presenting a wavelet enhanced appearance model using a bi-orthogonal CDF 9-7 wavelet basis. In [63], Larsen et al. presented a texture enhanced appearance model using Wedgelets. The Wedgelet representation was shown to be superior to wavelet representation at higher dimensionality reduction rates.

Statistical models of shape and appearance have also been used in the literature to simulate deformations in medical images and to generate ground truth data. These tasks are becoming important as simulated ground truth data can be useful for validating segmentation and registration algorithms. These models can also be used to artificially enlarge the size of training sets. Various methods have been proposed in the literature to achieve image synthesis. Traditionally, the simulation of the deformations in medical imaging is performed by using PCA-based methods. However, these methods are often

unable to capture realistic and complex deformations of the anatomical structure. Abboud et al. [64] proposed a method for facial expression synthesis by using PCA-based appearance models. In [65], Seals et al. proposed a method for enlarging the training set by using image morphing. In [66], Lee et al. presented a method for morphing multiple images. In [67], Hamarneh et al. presented a method for simulating the ground truth validation data via physically- and statistically-based warps. Rose and Taylor [68] proposed a statistical model for mammographic appearance by using steerable decomposition and PCA modeling. In [69], ElSafi et al. proposed a method for the statistical simulation of deformations by using Independent Component Analysis. In [70], Xue et al. presented a PCA-based method for simulating deformations in MR brain images.

2.6.3 Model-based segmentation methods

The segmentation of anatomical structures in medical images is a pre-request stage for various computer-assisted medical applications such as shape analysis, computer-aided diagnosis and surgical planning. Manual segmentation methods are both time-consuming and non-reproducible. Computer-assisted segmentation methods aim at reducing the time and cost of the error-prone manual segmentation process. In chapter six, we present a novel model-based vertebral segmentation method using local appearance profiles and a multi-scale shape prior. This section summarizes some of the related segmentation methods in the literature.

The work of Staib and Duncan [13] is one of the earliest publications to suggest using statistical prior models to guide the segmentation process. The authors proposed using elliptic Fourier descriptors to represent the object contour. The Fourier coefficients are trained by using a Gaussian distribution as the prior probability. Another milestone is the Active Shape Model (ASM) proposed by Cootes in [35]. Some of the basics of Active Shape Models (ASM) were presented in section 2.4. In the field of medical imaging, researchers now widely accept that the Active Shape model (ASM) suffers from a number of inherent limitations. First, being a PCA-based model, it fails to generalize well to unseen deformities of the anatomical structure. Second, the intensity-based appearance model in ASM inherently assumes that the object of interest has clearly identified boundaries, but this assumption is not necessarily true for medical images.

Several approaches have been proposed in the literature to address the challenges of prior-based segmentation methods. To improve the matching between the prior model and the unseen image, researchers have suggested using various image representations and features. Scott et al. [71] suggested using local features such as oriented gradients, corners and edge strengths. Van Gienneken [72] proposed a new local appearance profile using responses from a filterbank of Gaussian derivative functions. In [73], Li et al. suggested using steerable filters to extract the local edge features and used these features to construct the local appearance profile. In [74], De Bruijne and Nielsen proposed an image segmentation method using shape particle filtering. Seghers et al. [75] proposed an enhanced gray level local appearance model extracted from feature images. Zhu et al. [76] proposed a method to improve the search procedure of the Active Shape Model (ASM) by using Mixture of Gaussians for the grey-level profiles. Metaxas et al. [77] suggested a hybrid deformable model for segmentation and registration. In [78], Langs proposed the Active Feature Model (AFM), in which local texture descriptors are used. A number of researchers have also tried to develop more efficient shape prior models. In [79], Uzumcu et al. suggested using Independent Component Analysis (ICA) for modeling the shape variations in medical imaging. In [45, 80], Davatzikos et al. proposed a hierarchical active shape model using the wavelet transform and Principal Component Analysis. In [81], Koikkalainen, and Lotjonen presented a method for segmenting the atria in MR images by using combined PCA and ICA shape priors.

2.6.4 Segmentation and analysis of vertebral structures

In this thesis, we are particularly interested in the segmentation and quantification of human vertebrae in x-ray images. Accurate vertebrae segmentation is important for the proper assessment of various vertebral abnormalities. This section summarizes some of the work in the literature related to vertebral segmentation and shape analysis.

Over the past decade, several model-based segmentation methods have been proposed in the literature to extract both normal and pathological vertebral structures. In [56], Roberts et al. proposed a method for segmenting fractured vertebrae by using linked active appearance models. Howe et al. [82] presented a hierarchical method for segmenting cervical and lumbar vertebrae by using the Generalized Hough transform and the Active Appearance Model. The Generalized Hough transform is first used to estimate the initial pose. The Active Appearance Model (AAM) is then used to reach the final segmentation. The authors reported a segmentation accuracy of 4.35 mm for the lumbar spine. De

Bruijne et al. [74] reported a mean point-to-contour error of 1.4 mm for lumbar spine segmentation using shape particle filtering. This method is most useful in offline applications since it has a time-consuming pixel classification stage.

Statistical shape analysis refers to the ability to understand how anatomical structures tend to vary for various reasons such as the development of a certain pathology, aging, and surgery. Computer-aided diagnosis for vertebral abnormalities has been the focus of various research groups worldwide [83]. The common pathologies of interest include anterior osteophyte, disc space narrowing, inter-vertebral disc degeneration, and fractures.

The conventional methods for the diagnosis of vertebral deformities use a six-point representation for the vertebral body. Traditionally, most clinical practices have used this representation to assess various vertebral deformities. The six-point representation is usually used to derive measurements such as Anterior Height (AH), Middle Height (MH), and Posterior Height (PH). Based upon these heights, many methods have been proposed to quantify various vertebral deformities such as fractures [84, 85]. The major limitations of these methods are as follows: (i) they are subjective to the user, and usually fail to capture subtle shape changes, and (ii) the trade-off between specificity and sensitivity remains problematic.

Recent studies have attempted to find better shape representations for the analysis of vertebral structures. The National Library of Medicine (NLM) has been long interested in the vertebral classification of spine x-ray images for content-based image retrieval applications. Researchers at the engineering division of the NLM have implemented and evaluated several shape representations for the retrieval of spine x-rays by using anterior osteophyte pathology. Examples include Fourier descriptors, polygonal approximations, invariant moments, convex-hull based features, and partial shape matching using dynamic programming [86-89]. Some other researchers suggested using Principal Component Analysis (PCA) for the analysis and classification of vertebral deformities. Symth et al. [90] suggested using the Point Distribution Model (PDM) to represent vertebral shapes and employed a Mahalanobis distance classifier to identify vertebral fractures. Roberts et al. [91] suggested using Active Appearance Models for the segmentation and assessment of vertebral abnormalities. Promising results were reported. De Bruijne et al. [93, 94]

proposed a method for the classification of vertebral fractures in x-ray images by using pairwise conditional shape models trained on a set of healthy spines. Roberts et al. [92] proposed a method for vertebral classification using parameters from Active Shape Models and Active Appearance Models. A final comment regarding vertebral quantification is that none of the methods mentioned in this section is capable of analyzing multiple vertebrae simultaneously; in other words, no spatial information about the pathology is available.

3. Multi-scale Medical Image Analysis Tools: Wavelets and Beyond

3.1 Introduction

Advances in the field of computational harmonic analysis have been inspired partly by our understanding of the Human Visual System (HVS). Figure 3.1 shows a model of the low-level vision system as described in [95]. Vision researchers now believe that receptive field cells exhibit a localized response in both time and frequency. Various harmonic analysis tools attempt to mimic the behaviour of these cells. The work of Burt and Adelson [157] is among the earliest efforts in this field. The authors introduced the Laplacian Transform as a tool for multi-scale image analysis. In an independent effort, some mathematicians developed the wavelet transform, in which the basis functions are translations and dilations of a mother wavelet function. It is now accepted that a considerable amount of similarity exists between the response of the mammalian visual system and wavelet basis [95-97]. In [95], Olshausen and Field describe the receptive fields in the mammalian visual cortex as follows: “The receptive fields in the mammalian visual cortex can be characterized as being spatially localized, oriented, and band-pass comparable to the basis functions of wavelet transform.”

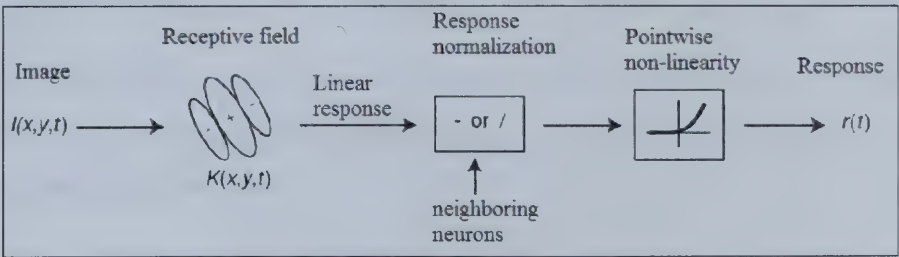


Figure 3.1 Model of the human low-level vision system [95]

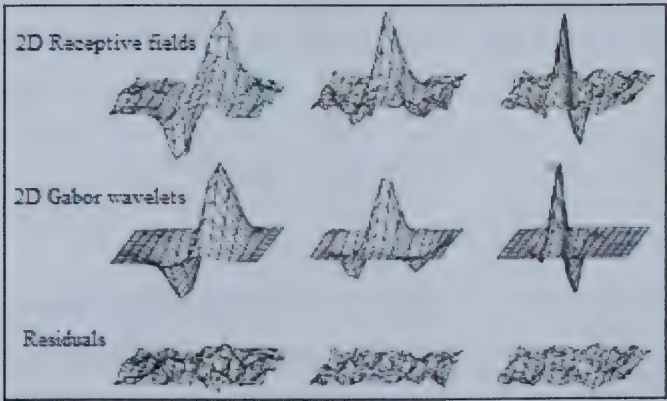


Figure 3.2 Similarity between receptive fields in mammalian primary visual cortex and Gabor wavelets [96].

Traditionally, wavelets have been extensively used for image analysis and interpretation tasks [98]. To overcome some of the limitations of traditional wavelets, a number of multi-scale analysis tools were introduced, including Non subsampled wavelets, Gabor filters, Steerable pyramids, Contourlet transform, and Non Subsampled Contourlet transform (NSCT). This chapter provides an overview of some of the recent advances in multi-scale image analysis tools. The chapter also discusses the suitability of various multi-scale tools for the analysis and quantification of medical images. Please refer to [97-99] for a more detailed theoretical background of the transforms.

3.2 Wavelet transform

3.2.1 Definition

Since their introduction, wavelets have become a popular tool for signal and image analysis in a wide range of applications including data compression, de-noising, and feature extraction. A wavelet transform decomposes a signal over a set of basis functions obtained by the scaling and translation of a mother wavelet function. A mother wavelet is a function that satisfies the admissibility condition:

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|^2} d\omega < \infty, \quad (3.1)$$

where $\Psi \in L^2(R)$, and $\Psi(\omega)$ is the Fourier transform of $\Psi(t)$. If $\Psi(\omega)$ has sufficient decay, the admissibility condition reduces to the constraint in which the mother wavelet function has zero average:

$$\int_{-\infty}^{\infty} \psi(t) dt = \psi(0) = 0. \quad (3.2)$$

A family of wavelet basis functions is obtained by scaling and translating $\Psi(t)$ by a and b , respectively:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right). \quad (3.3)$$

Two functions are needed to extract the low and high frequency components: the scaling and wavelet functions. Given a wavelet function $\Psi(t)$ and its companion scaling function $\phi(t)$, the following condition must be satisfied:

$$\phi(t) = \sqrt{2} \sum_b h(b) \phi(2t-b) \quad (3.4)$$

$$\psi(t) = \sqrt{2} \sum_b g(b) \phi(2t-b), \quad (3.5)$$

where $h(b)$ & $g(b)$ are related as follows:

$$g(b) = (-1)^b h(1-b). \quad (3.6)$$

The wavelet transform of a signal is defined as follows:

$$\mathbf{x}_{WT}(a, b) = \int_{-\infty}^{\infty} x(t) \psi_{a,b}(t) dt, \quad (3.7)$$

where $x \in L^2(R)$.

Multi-resolution Analysis (MRA)

By using two orthonormal sets of wavelet and scaling basis functions, the wavelet transform can be used to perform multi-resolution analysis. In the space of the finite real functions $L^2(R)$, we can define a sequence of resolutions such that signals at a scale smaller than 2^{-j} are suppressed at the j^{th} resolution level. Let V_j define the subspace of functions containing the signal information down to the 2^{-jth} scale. The multi-resolution representation is then obtained by decomposing the signal into a sequence of subspaces $V_J \dots \subset V_{-1} \subset V_0 \subset V_1 \dots \subset L^2(R)$, where

$$V_{j+1} = V_j \oplus W_j. \quad (3.8)$$

W_j is the detail space at the j^{th} resolution level and is orthogonal to the approximation space (V_j). The approximation space at a certain resolution can then be written as a sum of all subspaces, as shown in equation 3.9 [60]:

$$V_{j+1} = W_j \oplus V_j = W_j \oplus W_{j-1} \oplus V_{j-1} = \dots = W_j \oplus W_{j-1} \oplus W_{j-2} \dots \oplus W_{j-L} \oplus V_{j-L}. \quad (3.9)$$

The wavelet transform decomposes the signal into approximation and detail signals. A multi-resolution analysis (MRA) can then be achieved by iterating the decomposition process onto the approximation signal. The multi-resolution decomposition of a function is given by:

$$f(x) = \sum_{(i,j)} a_{M,i,j} \phi_{M,i,j} + \sum_{m=1}^M \sum_{(i,j)} \sum_{k=1}^3 d_{m,i,j}^k \psi_{m,i,j}^k, \quad (3.10)$$

where $a_{M,i,j} = \langle f, \phi_{M,i,j} \rangle$ and $d_{M,i,j}^k = \langle f, \psi_{M,i,j}^k \rangle$ are the approximation and wavelet detail coefficients, respectively. $\langle ., . \rangle$ is the inner product that orthogonally projects the function f onto the scaling or wavelet basis. M represents the highest decomposition level of the wavelet transform.

The wavelet decomposition of a two-dimensional signal is performed by applying the 1D wavelet transform onto the rows and the columns. Figure 3.3 illustrates a single decomposition step of a two-dimensional image. The process is repeated iteratively to obtain the multi-scale representation. Figure 3.4 illustrates the image decomposition using XY separable 2D wavelet filters.

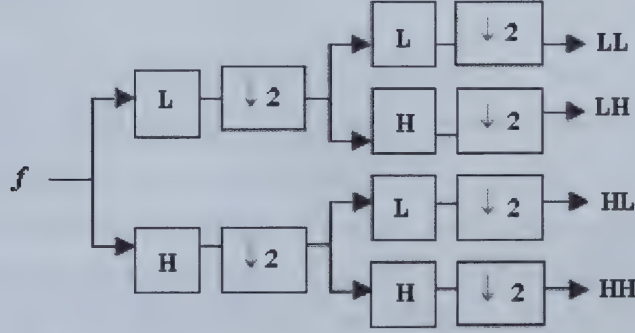


Figure 3.3 Wavelet decomposition of a two-dimensional signal

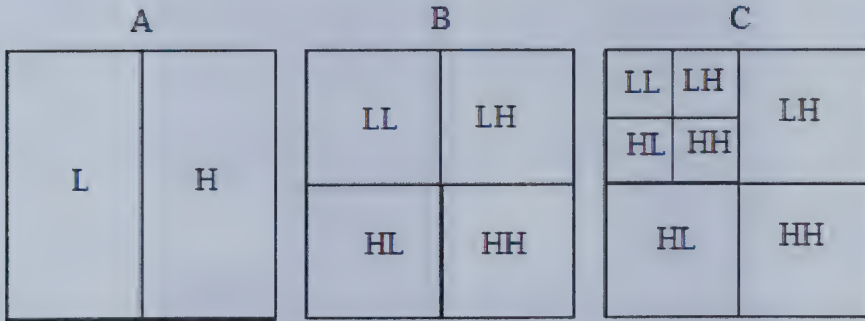


Figure 3.4 Image decomposition by using XY separable 2D wavelet filters

3.2.2 Wavelets in medical image analysis

Traditionally, wavelets have been extensively used in various image analysis and computer vision tasks. In the field of medical imaging, wavelets have proven useful in applications such as segmentation, registration, and image enhancement [114]. Their

usefulness is attributed to a number of appealing properties such as multi-scale representation and joint spatial-frequency localization. Over the past years, different wavelet families have been introduced in the literature. Each wavelet family possesses its own significance and drawbacks. Table 3.1 summarizes the properties of four well known wavelet families.

Table 3.1 Properties of four well known wavelet families

Wavelet\Property	Symmetry	Vanishing moments	Compact support	orthogonality
Haar	yes	1	yes	yes
Daubechies	far from	N	yes	yes
Coiflet	near from	2N	yes	yes
Bi-orthogonal	yes	N	yes	No

Despite their success, wavelets suffer from a number of limitations affecting their performance in medical imaging tasks. Some of these limitations are listed below:

- *Lack of orientation selectivity:* Wavelets suffer from poor orientation selectivity. Wavelet basis functions have only three permissible orientations. Thus, the wavelet transform is not capable of capturing rich orientation information about the structure of interest. Figure 3.5 illustrates the incomplete orientation selectivity of the wavelet transform.



Figure 3.5 Wavelet decomposition of an x-ray image of the lumbar spine

- *Lack of shift invariance:* Standard wavelet transform suffers from a lack of shift invariance due to the down-sampling step. Wavelet coefficients tend to change dramatically due to the translation of the input signal. This problem limits the capabilities of the transform as an image analysis and feature extraction tool.
- *Sub-band mixing and information shift:* Despite the computational advantages of orthogonal wavelets, they suffer from phase non-linearity and weight-shift property. The non-symmetry of orthogonal wavelets causes an information shift among different bands because of the projection onto biased basis functions. This information shift property limits the efficiency of the transform in localized image analysis tasks. Figure 3.6 illustrates the information shift property for Daubechies wavelets.

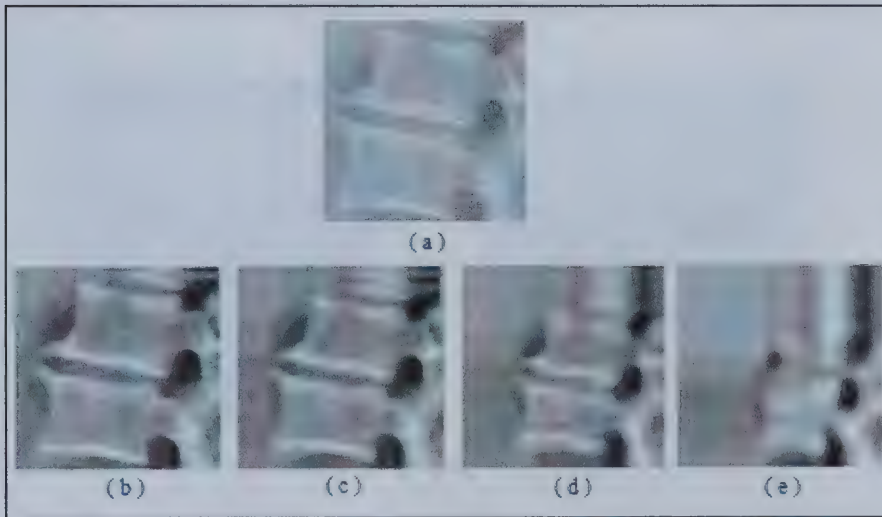


Figure 3.6 Information shift property for four-level orthogonal wavelet decomposition.

- *Response to higher order singularities:* In [99], Do and Vetterli showed that wavelets are optimal only in capturing point singularities. Wavelets are adapted only to abrupt changes in an image due to the XY separability of the basis functions. Thus, wavelets are not optimal in capturing line and curve singularities. A two-dimensional wavelet function “sees” a smooth contour as a number of point singularities. This feature decreases the efficiency of the transform in representing smooth contours and increases its sensitivity to noise. Medical images often exhibit higher-order singularities. Thus, the performance of

wavelets in medical image analysis is restricted by their ability to capture only horizontal, vertical and diagonal discontinuities. This problem is illustrated in Figure 3.7(a), where a large number of coefficients is needed to represent a smooth contour using wavelet bases. This problem obviously limits the performance of the wavelet transform as a feature extraction tool. Figure 3.7 (b, c) shows an example where a wavelet-based edge map fails to capture the noisy contours of the vertebral bodies in an x-ray image.

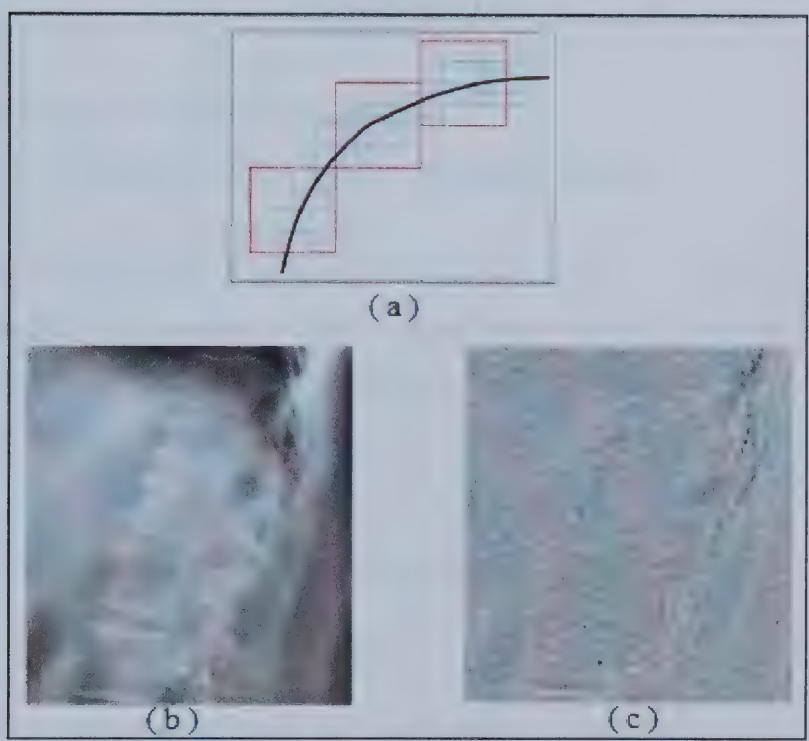


Figure 3.7 Illustration of the limitations of wavelets in representing contours. (a) Wavelet representation of a contour, (b) an x-ray image of lumbar spine with noisy contours, (c) wavelet-based edge map

3.3 Gabor Transform

3.3.1 Definition

In the seminal paper “Theory of Communications,” Denis Gabor introduced the concept of localization in time and frequency [100]. He argued that a useful decomposition basis function should be jointly localized in both time and frequency. The concept of joint localization aims at finding the minimal amount of information bounded by the uncertainty principle:

$$\Delta t \Delta f \geq \frac{1}{4\pi}, \quad (3.11)$$

where Δt : The time duration of the signal.

Δf : The bandwidth of the signal in the frequency domain.

The function which achieves the minimal boundary of the uncertainty principle is the result of modulating a harmonic oscillation of any frequency with a pulse in the form of a Gaussian distribution function [100]. A 1D Gabor function is given by

$$\psi(t) = e^{-\alpha^2 (t-t_o)^2} e^{j(2\pi f_o t + \varphi)}, \quad (3.12)$$

where α is the sharpness of the Gaussian function,

t_o is the centroid of the Gaussian,

f_o is the frequency of harmonic oscillations,

φ is the phase shift of oscillations.

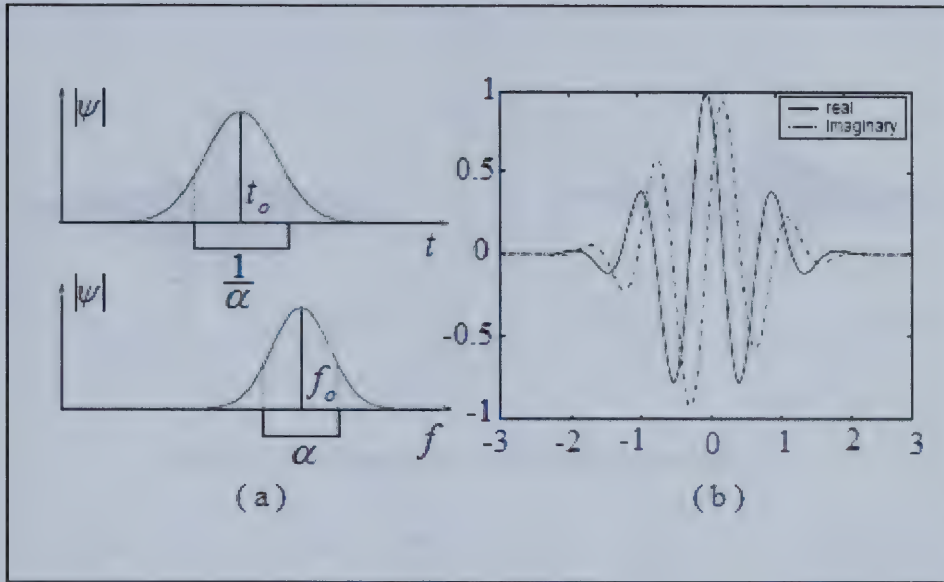


Figure 3.8 (a) One-Dimensional Gaussian function in time and frequency domain, (b) real and imaginary parts of Gaussian function

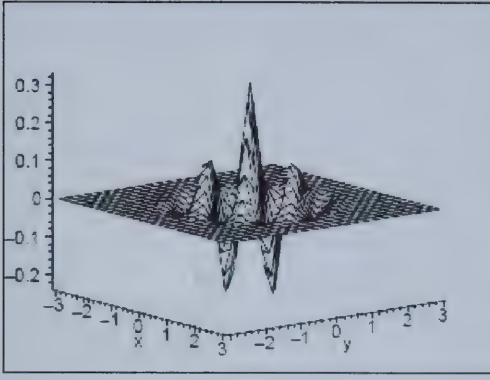
In [101], Daugman extended the 1D Gabor function to a 2D Gabor filter. A 2D Gabor function is defined as a complex sinusoidal plane wave modulated by an elliptic Gaussian probability distribution function. Mathematically, the 2D Gabor function is represented as follows:

$$\psi(x, y) = e^{-(\alpha^2 x'^2 + \beta^2 y'^2)} e^{j2\pi f_0 x'} \quad (3.13)$$

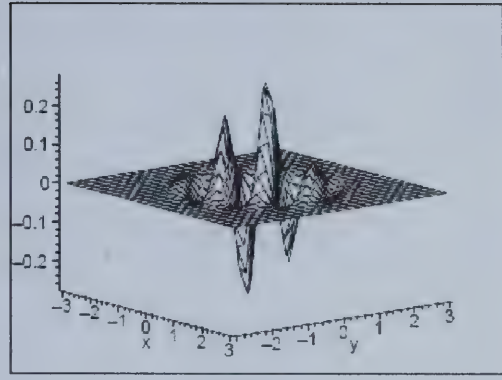
$$x' = x \cos \theta + y \sin \theta$$

$$y' = -x \sin \theta + y \cos \theta,$$

where f_0 is the frequency of the sinusoidal plane wave, θ is the anti-clockwise rotation of the Gaussian and the plane wave, α is the sharpness of the Gaussian along the axis parallel to the wave, and β is the sharpness along the axis perpendicular to the wave.



(a)



(b)

Figure 3.9 2-D Gabor filter (a) Real part; (b) Imaginary part

A set of Gabor filters can form a filterbank with each filter being characterized by parameters to specify the filter orientation, spatial frequency, and filter width. The filterbank is composed of filters at several orientations and frequencies. It has equal orientation and octave-frequency spacings:

$$\theta_k = \frac{k\pi}{n_\theta} \quad , \quad k = \{0, 1, \dots, n_\theta - 1\} \quad (3.14)$$

$$f_l = s^{-l} f_{\max} \quad , \quad l = \{0, 1, \dots, n_f - 1\} \quad , \quad (3.15)$$

where s (>1) is the ratio between two consecutive frequencies,

θ_k is the k^{th} orientation,

n_θ is the number of orientations,

f_l is the l^{th} frequency,

$f_{\max}$ is the maximal frequency.

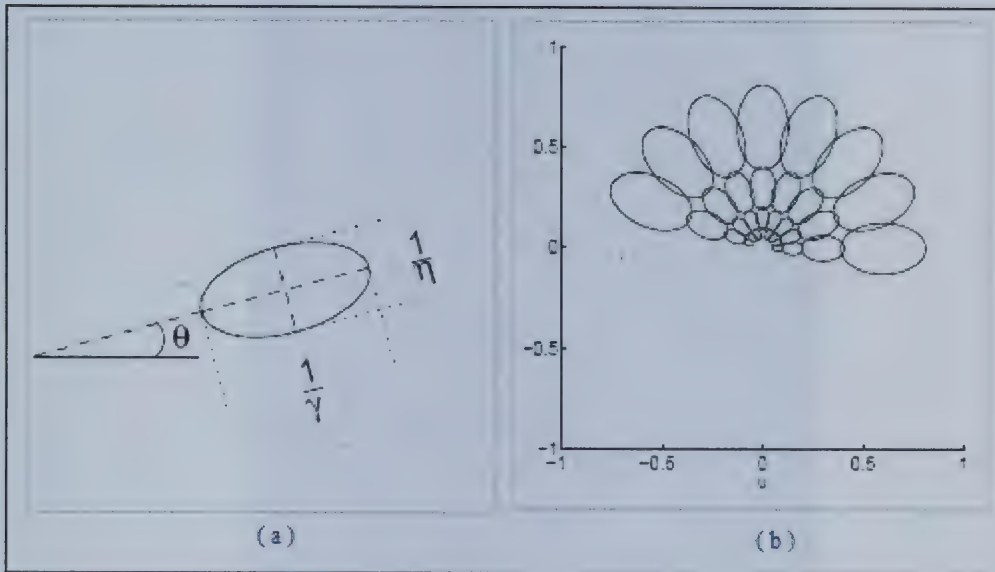


Figure 3.10 Gabor Filters. (a) Filter parameters in frequency domain; (b) bank of filters in frequency domain [102].

3.3.2 Gabor transform in medical image analysis

The Gabor transform has proven to be successful in performing various image analysis tasks. In comparison to wavelets, the Gabor transform offers better localization and orientation selectivity. In the field of medical imaging, Gabor filterbanks have been used for feature detection and representation tasks [103]. Various Gabor-based systems have been suggested for applications such as the enhancement, segmentation, and registration of medical images [102]. Figure 3.11 shows an example for the decomposition of a lumbar spine x-ray image by using a Gabor filter-bank.

Despite being able to encode rich information about local image structures, Gabor filters face a number of challenges. These challenges include the following:

- The selection of the filter parameters is a difficult task. A trade-off exists between the number of filters and the continuity of the feature space. In many imaging applications, the continuity of the feature space is a desirable property.
- In comparison to steerable pyramids [104], a Gabor filterbank does not have complete or approximate shift invariance.
- In comparison to Contourlets [99], Gabor transform allows only for a fixed number of orientations per scale.

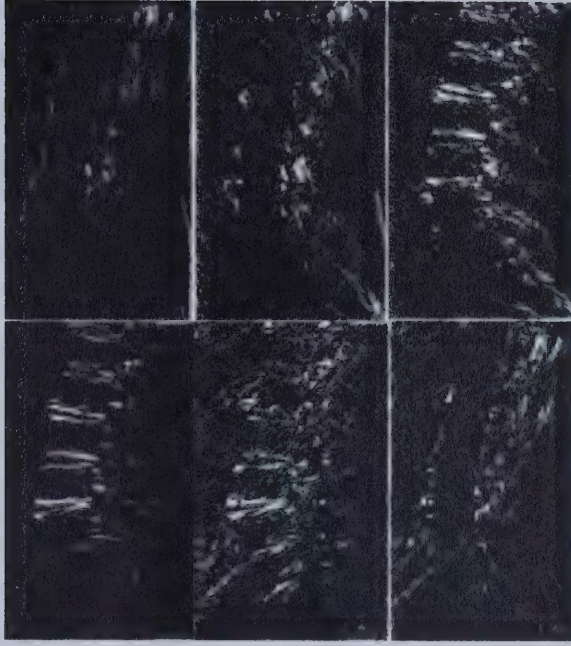


Figure 3.11 Analysis of lumbar spine x-ray by using Gabor Filterbank (six orientations at a fixed scale)

3.4 Steerable Pyramid

3.4.1 Definition

The steerable pyramid is a two-dimensional transform that is jointly-shiftable in orientation and position. Shiftability refers to the fact that the response at any parameter value (position or orientation) can be constructed as a linear combination from discrete values of the same parameter space [104]. Shiftability in position does not imply complete shift invariance. However, a shiftable transform exhibits “approximate” shift invariance where the power of its coefficients is invariant to translations of the input signal.

The basis functions of a steerable pyramid are the translations, dilations, and rotations of a single kernel. The transform is constructed as a recursive pyramid. Figure 3.12 shows a first-level image decomposition using a steerable pyramid. An input image is decomposed into low-pass and high-pass frequency bands. The low-pass frequency band is further decomposed into a number of oriented band-pass sub-bands $B_0, B_1, \dots, B_{K-1}$ and a lower low-pass sub-band L_1 . The low-pass sub-band is then subsampled by a factor of two, and the set of oriented band-pass filters are recursively applied.

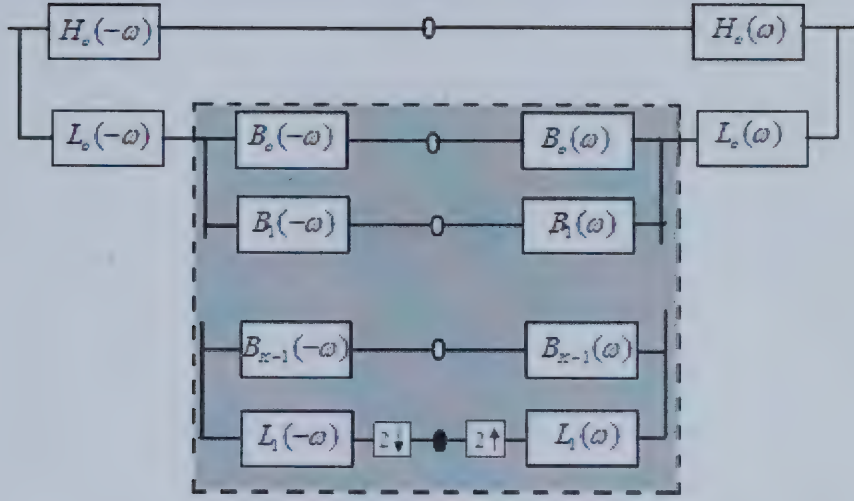


Figure 3.12 First level of Steerable pyramid decomposition

The first low-pass and high-pass filters are defined as

$$L_0(r, \theta) = L\left(\frac{r}{2}, \theta\right) / 2 \quad (3.16)$$

$$H_0(r, \theta) = H\left(\frac{r}{2}, \theta\right), \quad (3.17)$$

where r, θ are polar coordinates. $L(\cdot)$ and $H(\cdot)$ are raised cosine low-pass and high-pass filters:

$$L(r, \theta) = \begin{cases} 2, & r \leq \frac{\pi}{4} \\ 2 \cos\left(\frac{\pi}{2} \log_2\left(\frac{4r}{\pi}\right)\right), & \frac{\pi}{4} < r < \frac{\pi}{2} \\ 0, & r \leq \frac{\pi}{2} \end{cases} \quad (3.18)$$

The band-pass filter is a polar separable filter that is defined as follows:

$$B_k(r, \theta) = H(r)G_k(\theta) \quad (3.19)$$

$$H(r) = \begin{cases} \cos(\frac{\pi}{2} \log_2(\frac{2r}{\pi})), & \frac{\pi}{4} < r < \frac{\pi}{2} \\ 1, & r \geq \frac{\pi}{2} \\ 0, & r \leq \frac{\pi}{4} \end{cases}, \quad (3.20)$$

$$G_k(\theta) = \begin{cases} \alpha_k [\cos(\theta - \frac{\pi k}{K})]^{K-1}, & \left| \theta - \frac{\pi k}{K} \right| < \frac{\pi}{2} \\ 0, & \text{otherwise} \end{cases} \quad (3.21)$$

where $\alpha_K = 2^{K-1} \frac{(K-1)!}{\sqrt{K[2(K-1)]!}}$.

Figure 3.13 shows the spectral decomposition performed by the steerable pyramid. The basis functions are all dilations and rotations of a mother function except for the inner low-pass sub-band and the outer residual sub-band.

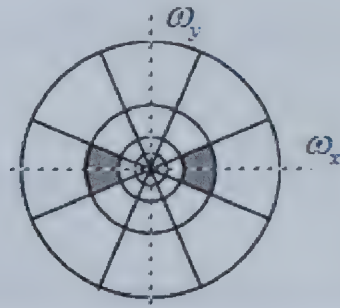


Figure 3.13 Illustration of the spectral decomposition performed by a Steerable pyramid.

3.4.2 Steerable pyramids in medical image analysis

- Steerable pyramids overcome many of the limitations in the wavelet transform. Steerable pyramids are jointly shiftable in orientation and position. This feature allows for better orientation selectivity over multiple scales, as shown in Figure 3.14.

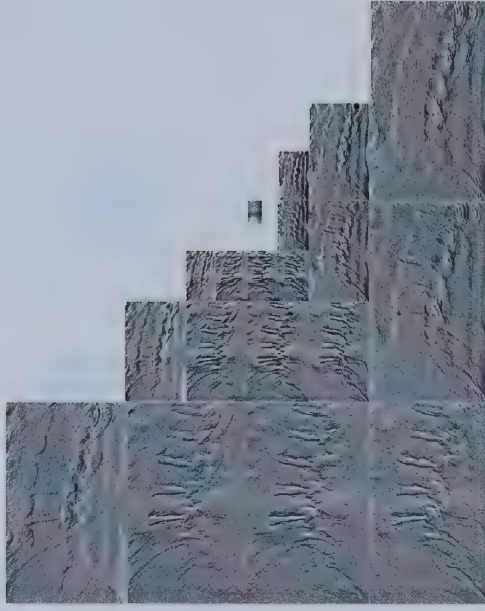


Figure 3.14 Analysis of lumbar spine x-ray by using Steerable pyramid

- In comparison to the Gabor transform, steerable pyramids offer better orientation analysis. Steerable pyramids allow for more flexible decomposition of the orientation space due to the steerability of the filters.
- Steerable pyramids are overcomplete. There are $\frac{4K}{3}$ times as many coefficients in the representation as in the original image. This feature is desirable in many image analysis and feature extraction tasks.
- In comparison to Contourlets [99], steerable pyramids can offer only a fixed angular resolution over all scales.

3.5 Contourlet Transform

3.5.1 Definition

In [99], Do et al. introduced the Contourlet transform as an efficient directional multi-scale tool for image representation. The transform was initially defined in the discrete domain. An efficient implementation for the transform was proposed by Do. et al. in [99]. The transform is implemented by using a pyramidal directional filter bank (PDFB). An image is decomposed into directional sub-bands at multiple scales. This decomposition is implemented by using a cascaded Laplacian pyramid, with a directional filter-bank applied at each scale. This decomposition is illustrated in Figure 3.15.

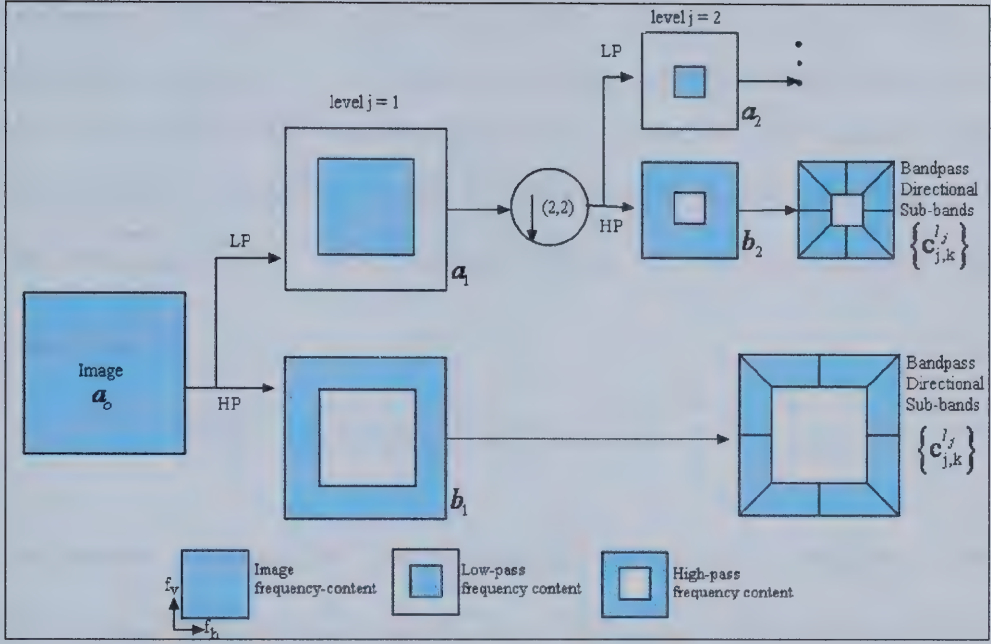


Figure 3.15 Contourlet decomposition by using ideal Pyramidal Directional Filterbank. The horizontal and vertical axes refer to the horizontal and vertical frequencies of an image. The coloured region shows the frequency components present in the image. The white region refers to the frequency band filtered out.

The Laplacian pyramid captures the point singularities, and the subsequent directional filtering captures the linear singularities. Multi-scale decomposition for the input image (a_0) is performed by using the Laplacian pyramid. The outcome is a set of J band-pass images and a low-pass image, as in equation (3.22):

$$LP_J(a_0) = \begin{cases} b_j, j = 1, 2, \dots, J \\ a_1 \end{cases}, \quad (3.22)$$

where LP_J is an operator for the Laplacian pyramid with J levels,

a_0 is the input image,

$\{b_j\}$ is the set of band-pass images arranged in a fine-to-coarse order,

a_j is a low-pass coarse image at the J^{th} resolution level.

Next, a multi-directional decomposition is achieved by applying an iterated directional filterbank (DFB) for each band-pass image (b_j) [99,105]. Do et al. constructed a 2-D

directional filterbank (DFB) that can be maximally decimated while preserving perfect reconstruction. This task is efficiently achieved through an l -level binary tree decomposition resulting in 2^l sub-bands with wedge-shaped frequency partitioning. The l -level tree-structured DFB can be equivalently viewed as a 2^l parallel channel filterbank with equivalent filters. At the j^{th} resolution level, the band-pass image (b_j) is further decomposed by an l_j -level directional filterbank (DFB). The outcome is a set of response images $\{c_{j,k}^{l_j}\}$:

$$DFB_{l_j}(b_j) = \{c_{j,k}^{l_j}\}, k = 0, 1, \dots, 2^{l_j} - 1 \quad (3.23)$$

In the frequency domain, the Laplacian pyramid (LP) creates a sequence of nested approximation subspaces, $\dots V_2 \subset V_1 \subset V_0 \dots$, where V_j is defined on a uniform grid with intervals $2^j \times 2^j$. This pyramid also creates the detail subspaces, $\dots W_2 \subset W_1 \subset W_0 \dots$, which represent the difference images in the Laplacian pyramid that hold the details necessary to increase the resolution between any two consecutive approximation subspaces. The DFB is then applied onto the detail subspace W_j . Figure 3.16 depicts the division of the frequency space by using the Contourlet transform. V_j is the approximation subspace. W_j is the detail subspace. The index k runs from 0 to $2^l - 1$ directions [99].

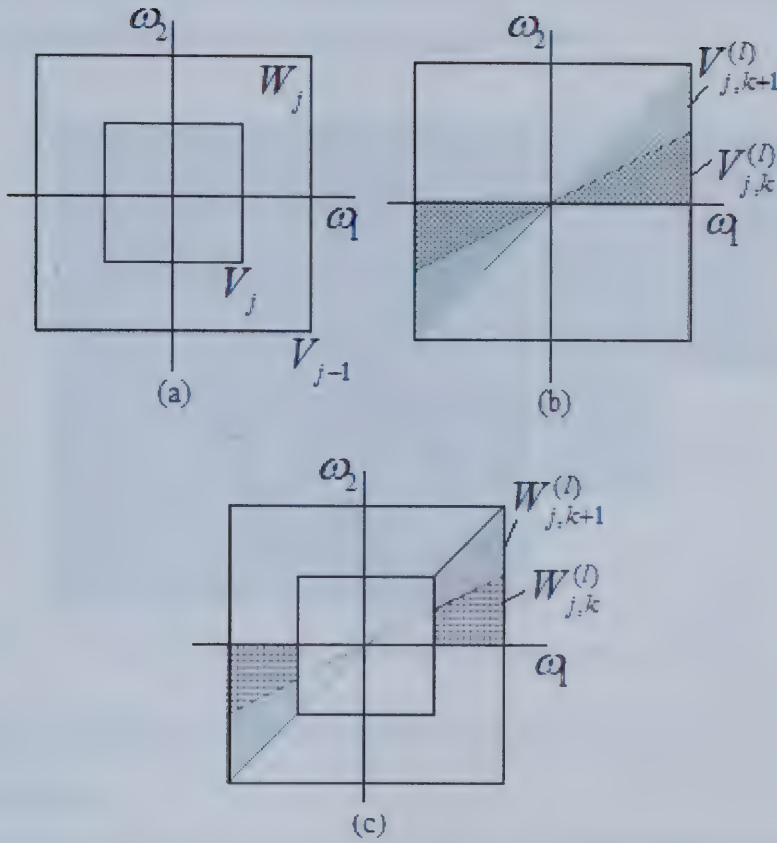


Figure 3.16 Division of frequency spectrum by using Contourlet transform.(a) Subspaces generated by LP, (b) Subspaces generated by DFB, (c) Complete division generated by PDFB .

3.5.2 Contourlets in medical image analysis

The Contourlet transform has a number of interesting properties that make it an appealing alternative for medical image analysis. The Contourlet transform provides an efficient sparse representation for two-dimensional signals. It possesses a number of desired features such as multi-resolution representation and a high degree of directionality, time-frequency localization, near-critical sampling, and anisotropy. These features allow Contourlets to deal effectively with smooth contours in medical images. The Contourlet transform is capable of performing flexible directional decomposition per scale, while keeping a near-critical sampling. The Contourlet transform is also computationally efficient as it is constructed by using iterated filterbanks.

Owing to these properties, the Contourlet transform has the potential to outperform the wavelet transform in various image analysis tasks. Nevertheless, it still suffers from a

lack of shift invariance due to the down-sampling step. Shift invariance is a desirable feature in tasks such as feature extraction and edge detection.

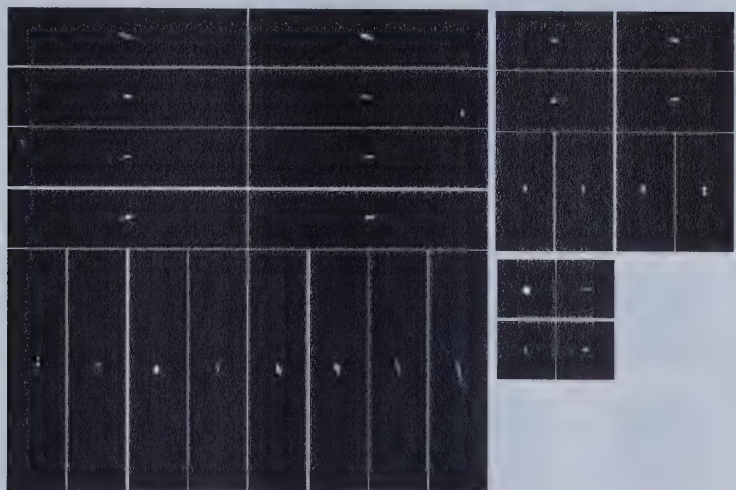


Figure 3.17 Example of basis functions of contourlet transform

3.6 Non Subsampled Contourlet transform (NSCT)

3.6.1 Definition

In [106], Do. et al. introduced the Non Subsampled Contourlet Transform (NSCT) as the shift invariant version of the Contourlet transform. The NSCT is a shift invariant multi-scale multi-directional image representation tool. The NSCT is constructed by using non subsampled pyramids and non subsampled Directional Filter-Banks (NS-DFB). The multi-scale decomposition is achieved by using the non subsampled pyramid. The non-subsampled Directional Filter Bank (NS-DFB) is responsible for the multi-directional analysis per scale.

Figure 3.18 shows a block diagram for the Non Subsampled Contourlet Transform (NSCT). The non subsampled pyramid filters the image into low-pass and high-pass bands. The high-pass band is further decomposed by using the non subsampled Directional Filterbank (NS-DFB) to obtain several directional sub-bands. The process is iterated onto the low-pass sub-band. Figure 3.18(b) shows the resulting partitioning of the frequency domain.

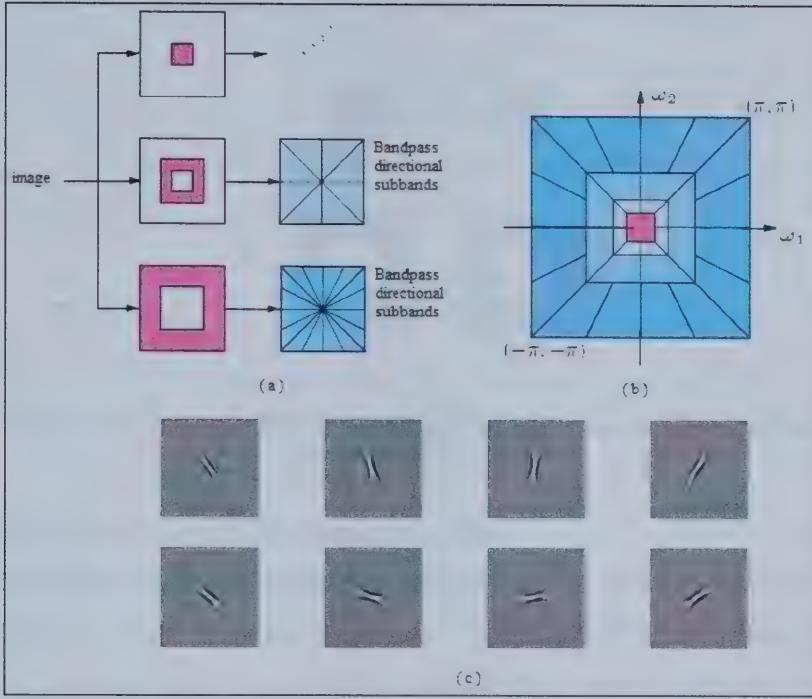


Figure 3.18 Non Sub-sampled contourlet transform. (a) NSCT filterbank. (b) Idealized frequency partitioning, (c) NSCT basis functions.

Non Subsampled Pyramid

A non-subsampled pyramid is used to achieve multi-scale decomposition. Shift invariance is preserved due to the non sub-sampled decomposition step. The building block of the non subsampled pyramid is a two-channel non subsampled 2D filterbank that satisfies:

$$H_o(z)G_o(z) + H_1(z)G_1(z) = 1, \quad (3.24)$$

where $H_o(z)$ and $H_1(z)$ are the low-pass and high-pass analysis filters. $G_o(z)$, $G_1(z)$ are the low-pass and high-pass synthesis filters. The pyramid is constructed by iterating the non subsampled filterbank.

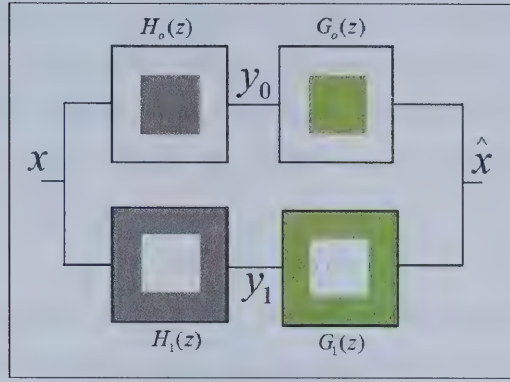


Figure 3.19 Non subsampled pyramid decomposition. Frequency response of the building block of the non subsampled pyramid.

Non subsampled Directional Filter Bank (NS-DFB)

A Non subsampled Directional Filter-bank (NS-DFB) is used to achieve multi-directional analysis at every scale. The building block of the NS-DFB is a two-channel non subsampled 2D filterbank. Figure 3.20 shows the frequency response of the building block of the non subsampled Directional Filterbank (NS-DFB). Finer directional decomposition is achieved by iterating the process.

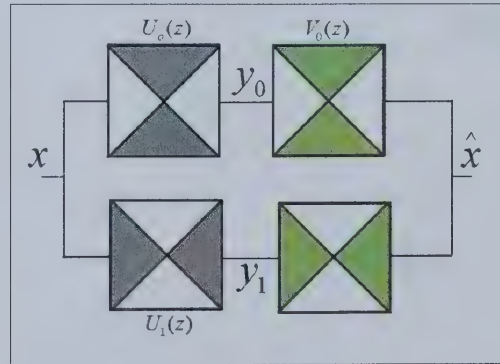


Figure 3.20 Frequency response of the building block of the Non Subsampled Directional Fiterbank.

3.6.2 Non Subsampled Contourlet Transform in medical image analysis

Medical images are often examined to extract features and geometric information about underlying anatomical structures. The Non Subsampled Contourlet Transform (NSCT) exhibits a number of significant properties that make it suitable for the analysis and interpretation of medical images.

- The NSCT shares a number of interesting properties with Wavelets, such as the multi-scale and locality of the basis functions. In addition to these properties, the NSCT exhibits non-isotropic basis functions which allow for better orientation selectivity over scales. Thus, the NSCT can better capture higher order discontinuities in an image.
- In comparison to Gabor filters and steerable pyramids, the NSCT allows for variable angular resolutions over the scales. Thus, more powerful orientation analysis can be achieved by using the NSCT. The NSCT is also more computationally appealing than Gabor filters.
- In comparison to the Contourlet transform, the NSCT is shift invariant. This feature means that a pixel in the original image domain will maintain its spatial position in all sub-bands of the transform. Thus, we can extract more discriminatory and rich information about the structure under investigation. This benefit implies that the NSCT would be more suitable for applications involving feature extraction and object-tracking tasks.

4. Multi-scale Shape Model Using Best Basis Selection Framework and Independent Component Analysis

4.1 Introduction

In biomedical applications, statistical shape models try to faithfully represent the full range of biological variations in anatomical structures. These models are often used in a wide range of applications such as segmentation, registration, and surgical planning. We are particularly interested in applications to support medical practices related to the imaging of the human spine. Medical experts often examine hundreds of spine radiographs to determine the existence of various vertebral pathologies. Common pathologies of interest include anterior osteophyte, disc space narrowing, and fractures. Figure 4.1 shows examples of x-ray images of the human lumbar spine. Figure 4.2 shows two vertebrae with localized osteophyte pathology. The accurate modeling of anatomical structures is a challenging task due to shape complexity and the wide range of possible local and global variations. In this chapter, we attempt to tackle some of the existing challenges in shape modelling. In particular, we focus on the following goals:

- Finding a proper shape representation that accurately characterizes the anatomical structure of interest.
- Constructing an efficient statistical model that captures both local and global pathology-related deformations.



Figure 4.1 Examples of x-ray images of the human lumbar spine [108]

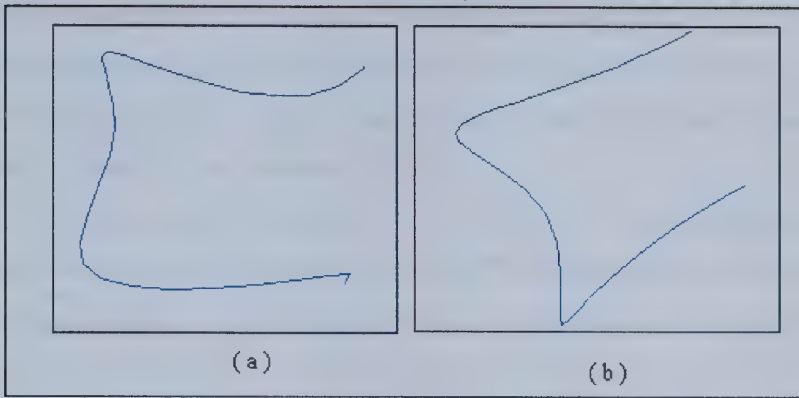


Figure 4.2 Two vertebrae with anterior osteophytes. (a) Lumbar vertebra with upper left-hand corner osteophyte, (b) lumbar vertebra with upper and lower left-hand corner osteophyte.

Within this context, this chapter presents a novel multi-scale statistical shape model using the concepts of sparsity, dimension reduction, and statistical independence to extract multi-scale modes of deformations. Within a best basis selection framework, the proposed model benefits from the multi-scale and sparsifying nature of wavelet packets and the ability of Independent Component Analysis (ICA) to capture higher-order statistics. This advantage allows us to construct a localized, non-uniform shape model with clinically relevant modes of deformations. The rest of the chapter is organized as follows: section 4.2 discusses some of the limitations of Principal Component Analysis (PCA) as a shape representation and modeling tool. Section 4.3 provides a description of the dataset used in our experiments. Section 4.4 describes shape modeling as a basis selection problem. Details of the proposed multi-scale shape model are presented in section 4.5. Experimental results are given in section 4.6. Discussions and conclusions are presented in sections 4.7 and 4.8, respectively.

4.2 Limitations of PCA-based statistical modeling

Principal Component Analysis (PCA) is a classical technique that is useful mainly for data reduction. Given a set of measurements, PCA tries to reduce the redundancy in the representation through decorrelation. In the original formulation of the Active Shape Model [35, 36], Cootes et al. described a Point Distribution Model (PDM) that captures the shape variations over a training set by using Principal Component Analysis (PCA). Details of the Active Shape Model were presented in section 2.4. Within the context of statistical modeling, PCA has the advantage of producing a compact representation. However, recent studies have indicated that PCA-based models suffer from a number of

inherent limitations that restrict the overall model performance [40, 49, 107], particularly with complex anatomical structures. These limitations include the following:

1. A PCA-based shape model assumes that the data are normally distributed, but this assumption is not necessarily true.
2. A PCA-based model represents the shapes as a linear combination of the eigenvectors plus an average shape. Using the eigenvectors as the basis functions limits the ability of the model to capture local variations. A change in the weight of one vector affects the entire shape. Thus, a PCA-based model tends to capture global variations more often than local variations. Figure 4.3 shows the first mode of variation for the Point Distribution Model (PDM) of the lumbar spine. Clearly, changing the weight of the mode globally affects the overall shape.

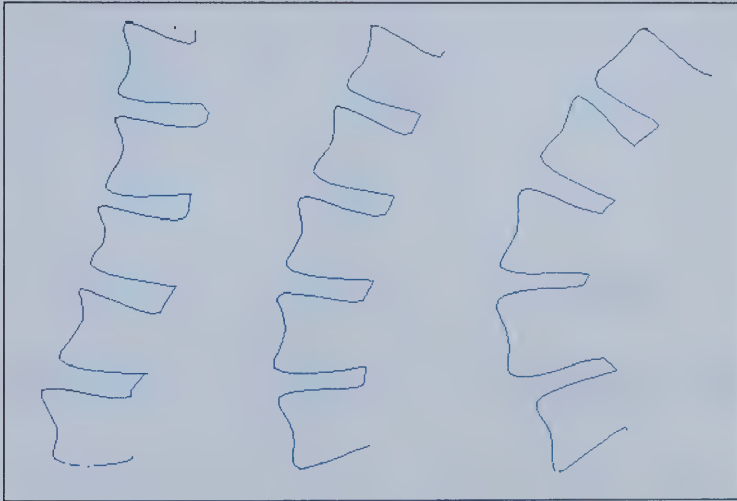


Figure 4.3 First mode of variation of a PCA-based model of lumbar spine

3. When the size of the training set is much smaller than the dimension of the shape space, then the sample covariance matrix estimated by using PCA becomes a poor estimator of the true covariance matrix. This limitation increases the model's sensitivity to noise and outliers.
4. The degree of freedom of a PCA-based model is limited by the number of samples in the training set. For a training set of $K-1$ shapes, each with M degrees of freedom and $M \gg K-1$, a PCA-based model will capture only up to $(K-1)$ modes of variations. Thus, with few training samples available, the model is not able to generalize well to valid shapes that are not present in the training set.

Figure 4.4 shows examples where a PCA-based model of the lumbar spine fails to capture localized deformations.

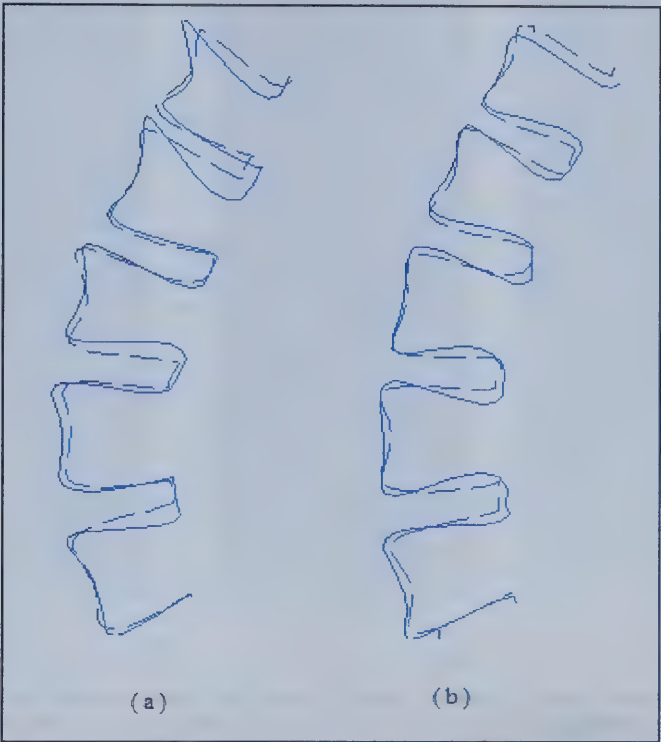


Figure 4.4 Examples for failure of PCA-based model to capture localized vertebrae deformations. The solid lines are the actual shapes, and the dotted lines are the predicted shapes.

In section 2.6.1, we discussed several approaches suggested in the literature to overcome the limitations of PCA-based shape models. The work presented in this chapter builds upon and extends these research efforts.

4.3 Data Description

The proposed multi-scale shape model presented in this chapter is evaluated by using a dataset of 40 outlined contours of five vertebrae from x-ray images. The spine x-ray dataset was provided by the National Library of Medicine [108]. For each x-ray image, the lumbar vertebrae L1-L5 are annotated and labelled as normal or abnormal (i.e., with osteophyte pathology) by experienced radiologists. Figure 4.5 shows how a dataset of lumbar spine shapes is built. By fitting a B-spline through the selected boundary in each image, we obtain a smooth outlined contour of the lumbar spine.

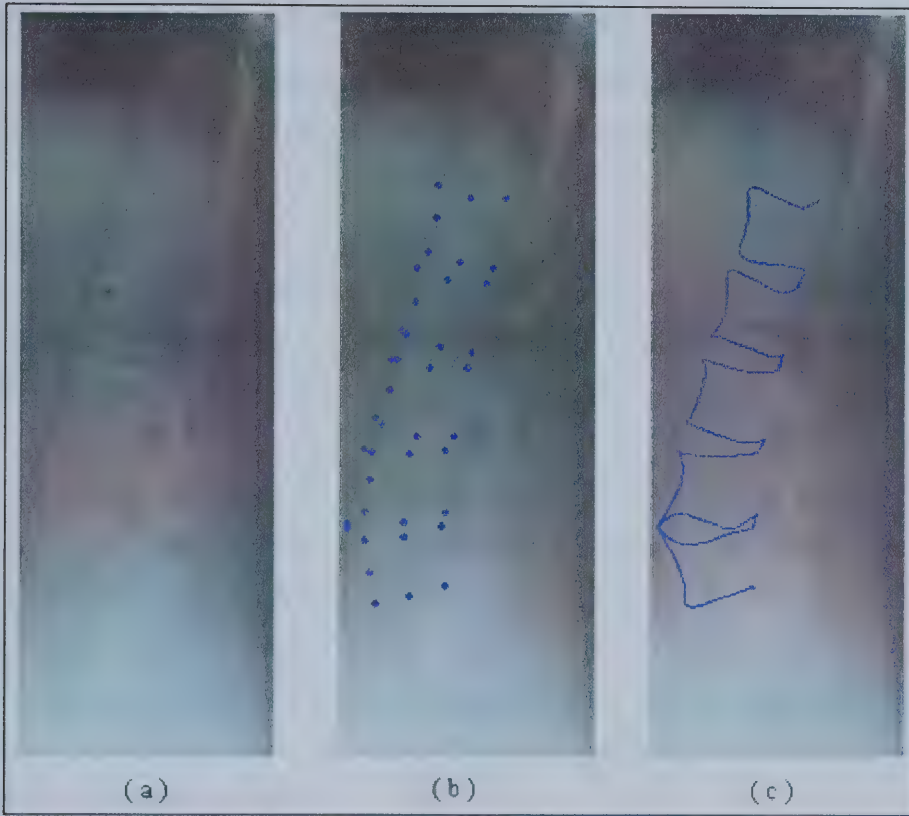


Figure 4.5 Building training dataset for shapes of lumbar spine. (a) Sample x-ray image of lumbar spine, (b) Boundary points selected by an expert, (c) Interpolated shape of lumbar spine

4.4 Shape modeling as a basis selection problem

It is now widely accepted that the PCA basis is not well suited for modeling complex anatomical shapes, which normally exhibit localized deformations [49, 52, 54, and 111]. Finding a suitable basis set for a particular class of signals is an important task in a wide variety of fields. In the context of modeling anatomical structures, desirable properties in the modeling basis include the locality, interpretability, and compactness of the representation. At the core of the model presented in this chapter are the concepts of multi-scale representation, sparsity, and statistical independence.

4.4.1 Multi-scale analysis

The concept of multi-scale analysis is supported by our biological vision functions. Over the past years, wavelets have been successively used as the predominant multi-scale representation tool. The wavelet transform offers an elegant way of scale-frequency decomposition by using localized basis functions. More details about wavelets were presented in chapter three. Multi-scale analysis involves decomposing the function into a

sequence of subspaces $V_J \dots \subset V_{-1} \subset V_0 \subset V_1 \dots \subset L^2(R)$ and then studying the signal response in each of these subspaces. The proposed shape model employs the Wavelet Packets basis to obtain a multi-scale representation of the shapes in the training set.

4.4.2 Sparsity

Recent studies have emphasized the role of signal sparsity in facilitating data interpretation. Sparse coding has proven to be useful in a diverse range of problems including biomedical inverse problems, spectral estimation, and data compression [109, 110].

The sparsity of a representation means that only few coefficients are significantly active, while most of the other components are close to zero. Most signals can be projected onto a sparse space by using a suitable sparsifying basis set, which can be chosen as a fixed set of functions (e.g., the Discrete Cosine Transform and the Wavelet transform). Another alternative is to search for a best-basis set in an overcomplete dictionary of basis functions (e.g., wavelet packets).

For a vector $x \in R^n$, the true sparsity measure is the quasi-norm L^0 , which is defined as

$$\|x\|_0 = \#\{i \in [1, n] : x_i \neq 0\}. \quad (4.1)$$

However, due to the instability of the true sparsity measure, it is usually replaced by the L^p norm, which is defined as

$$\|x\|_p = \left\{ \sum_{i=1}^N |x_i|^p \right\}^{\frac{1}{p}}, \quad 0 < p < 1. \quad (4.2)$$

4.4.3 Statistical independence

The statistical independence constraint on a basis set means that the basis functions do not interfere with each other. This fact can be explained in statistical terms as follows:

Let $p_1(x_1)$ and $p_2(x_2)$ denote the probability densities of two basis vectors x_1 and x_2 , respectively. Let $p(x)$ denotes the joint probability. x_1 and x_2 are said to be statistically independent if and only if the joint distribution function can be expressed as:

(4.3)

$$p(x) = p_1(x_1)p_2(x_2).$$

The statistical independence of x_1 and x_2 implies that they are uncorrelated:

$$COV(x_1, x_2) = E\{(x_1 - \mu_{x_1})(x_2 - \mu_{x_2})\} = E(x_1)E(x_2) - \mu_{x_1}\mu_{x_2} = 0. \quad (4.4)$$

While covariance captures the second-order dependences between x_1 and x_2 , statistical independence tries to capture higher-order statistics as well.

4.5 Methodology

This section outlines the steps for constructing the proposed multi-scale shape model by using best-based selection and Independent Component Analysis. Throughout this section, the proposed model is referred to as the Wavelet-ICA shape model.

We start with N-training samples in the form of K corresponding landmarks. The set of training samples are first aligned by using the Procrustes alignment method. Let the n^{th} shape vector be represented as

$$X_n = [x_{n1} \dots x_{nK}, y_{n1} \dots y_{nK}]^T, \quad (4.5)$$

where x and y are the x - and y -coordinates.

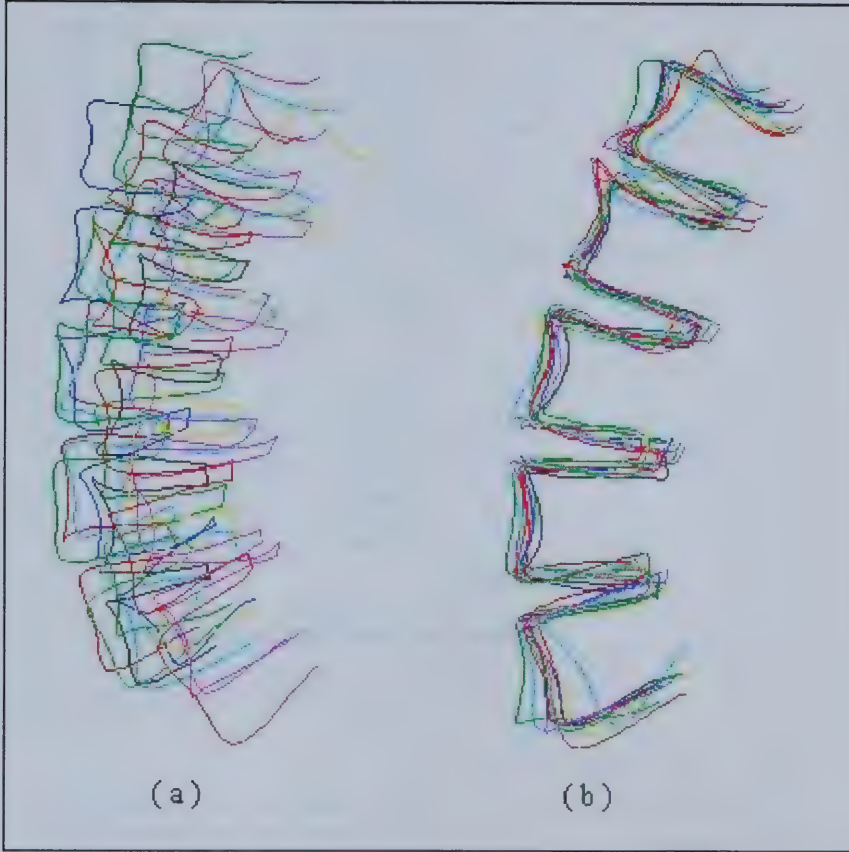


Figure 4.6 Shape alignment using Procrustes method (a) unaligned shapes of lumbar spine, (b) aligned shapes of lumbar spine.

Figure 4.7 shows the block diagram of the proposed Wavelet-ICA shape model. Given the training set of aligned shapes, a Multi-scale sparse representation is obtained by projecting the shapes onto a joint best-basis set selected from an overcomplete dictionary of wavelet packets. An entropy-based cost function is used for the best-basis selection. Next, we benefit from the approximating and sparsifying nature of wavelet packets to achieve efficient feature selection. The outcome is a reduced set of wavelet packet coefficients with high statistical significance. A spectral clustering step is next applied to extract bands of correlated coefficients. The statistical properties of the wavelet packet coefficients in each sub-band are then modelled by using Independent Component Analysis (ICA).

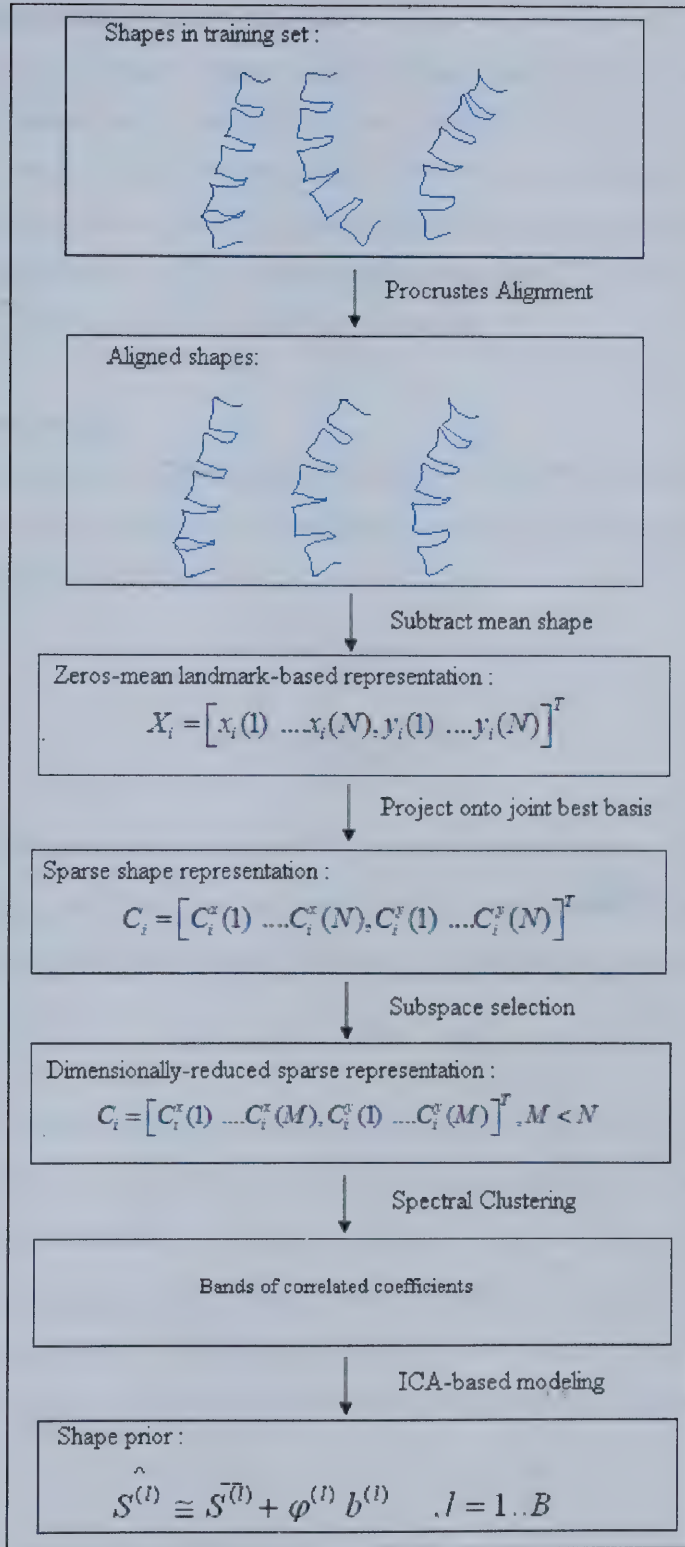


Figure 4.7 Block diagram of the proposed multi-scale Wavelet-ICA shape model

4.5.1 Multi-scale shape representation

Within the context of medical imaging, statistical shape models are required to have good statistical performance and relevant clinical interpretability. The modeling of anatomical structures is closely tied to the problem of efficient representation. Thus, prior to statistical modeling, we first find an optimal multi-scale representation of the anatomical structures in the training set. This representation is then used to construct a shape model with multi-scale and interpretable patterns of variations.

4.5.1.1 Wavelet Packets

Over the past years, wavelets have been successively used as the predominant multi-scale representation tool [97]. Wavelet packets represent a family of orthonormal bases in $L^2(\mathbb{R})$. The wavelet packet is defined as [112]

$$\psi^{2n}(t) = \sum_k h_k \psi^n(2t - k) \quad (4.7)$$

$$\psi^{2n+1}(t) = \sum_k g_k \psi^n(2t - k) \quad , \quad (4.6)$$

where $\psi^0(t)$ is the scaling function, and $\psi^1(t)$ is the wavelet function. $\{h_n\}, \{g_n\}$ are the low-pass and high-pass filters associated with the wavelet transform. A wavelet packet basis is any orthonormal basis selected from the set of functions in equation 4.8:

$$\psi_{k,i,j} = 2^{k/2} \psi_i(2^k t - j) \quad , \quad (4.8)$$

where k is the scale parameter, K is the maximum scale, $i = 0, \dots, 2^k - 1$ is the frequency index at the k^{th} scale. $j = 0, \dots, 2^{K-k} - 1$ is the translation index at the k^{th} scale, and 2^{K_0} is the length of the input signal.

The wavelet packet library can be organized in a binary tree structure, with each node representing the projection of the signal into a certain sub-space, as shown in Figure 4.8. A p -level wavelet packet tree divides the frequency domain into 2^p bands or sub-spaces.

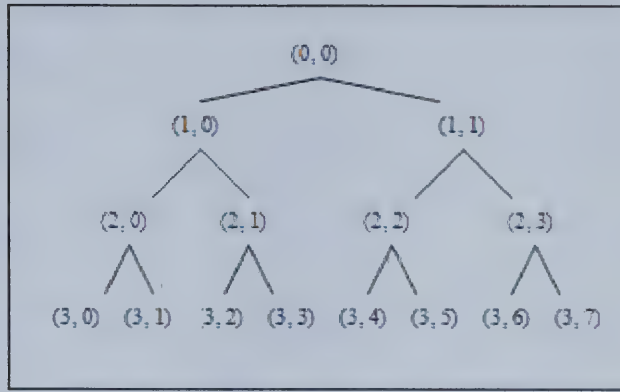


Figure 4.8 The tree of wavelet packet decomposition. Each doublet (k,i) represents the depth(k), and the number (i) of the nodes.

Let $\{\psi_{k,i}\}$ be the collection of all the wavelet packets of the translations at node (k,i) .

The collection of coefficients corresponding to the disjoint cover of the frequency domain will then form an orthonormal basis. Thus, for k levels of decomposition, up to 2^{2^k} bases are present.

This flexibility offered by wavelet packets is exploited to construct a multi-scale localized statistical model. Figure 4.9 shows an example of vertebral shape decomposition using wavelet packets. Due to the local support of the basis functions, they tend to capture different characteristics of the lumbar spine, ranging from the global aspects of the shapes to more localized deformations.

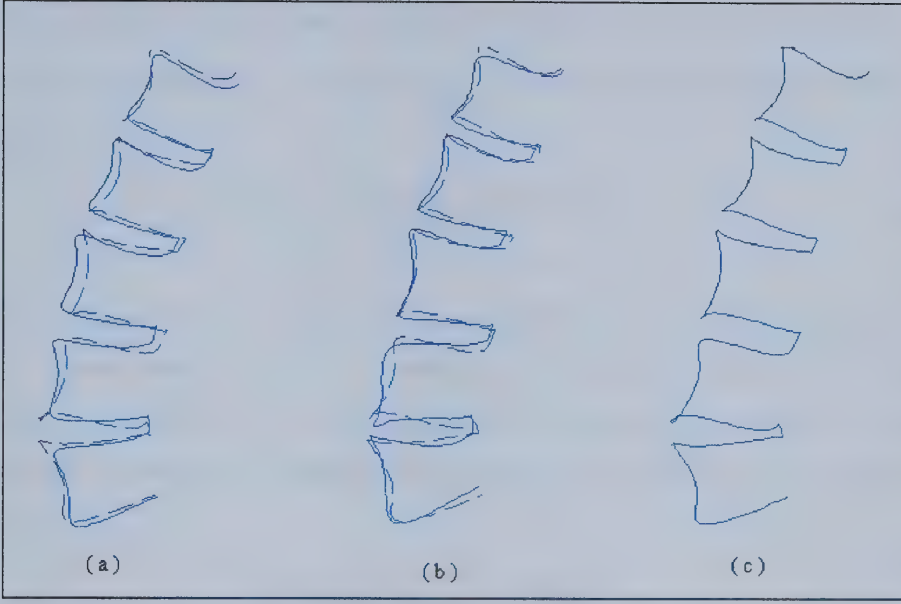


Figure 4.9 Example of shape decomposition using wavelet packets. (a) Mean shape, (b) Mean shape plus variations in low frequency bands, (c) Mean shape plus variations from all frequency bands. (The dotted shape is the original shape).

4.5.1.2 Optimal representation using best basis selection

The multi-scale representation is achieved by searching for the joint best-basis in an overcomplete wavelet packet dictionary. Let the shape vector be represented as

$$X = [x_1, \dots, x_N, y_1, \dots, y_N]^T, \quad (4.9)$$

where x and y are x - and y -coordinates.

We use wavelet packets to find an optimal representation for the shapes in the training set. For a p -level wavelet packet tree, there are 2^{2^p} possible orthonormal bases. Each basis corresponds to a specific decomposition for the frequency space. Given a library of wavelet packet basis, we search for the basis that best represents the shapes in the training set.

Let D denotes a wavelet packet dictionary with all 2^{2^p} possible orthonormal bases $\{\varphi_i\}$. The projection of the shape vector X onto any basis φ is then given by

$$C = \varphi^T X. \quad (4.10)$$

Let $M(\varphi^T X)$ denote the cost function for projecting the shape X onto the basis φ . The best basis in the dictionary D is then the one for which the cost function, $M(\varphi^T X)$, is minimized:

$$\varphi_{best} = \arg \min_{\varphi \in D} (M(\varphi^T X)) . \quad (4.11)$$

Coifman and Wickehauser [115] proposed a fast bottom-up search algorithm that finds the best orthonormal basis from a wavelet packet library according to an additive information cost function. In our model, we use the Shannon entropy as the cost function for selecting the best basis. The Shannon entropy for a wavelet packet node (k, i) is defined as

$$H(k, i) = -\sum p_j \log(p_j) , \quad (4.12)$$

Where $p_j = \frac{\|C_{k,i}(j)\|^2}{\|X\|^2}$, $C_{k,i}$ are the wavelet packet coefficients at the node (k, i) , X is

the observed shape signal, k is the scale parameter of the wavelet packet tree, and i is the frequency index at the k^{th} scale. j is the translation index at the k^{th} scale and i^{th} frequency index.

Using the best-basis selection framework, we obtain a multi-scale sparse representation for the anatomical shape of interest. In statistical terms, this result means that the distribution of the coefficients is highly peaked, as illustrated in Figure 4.10

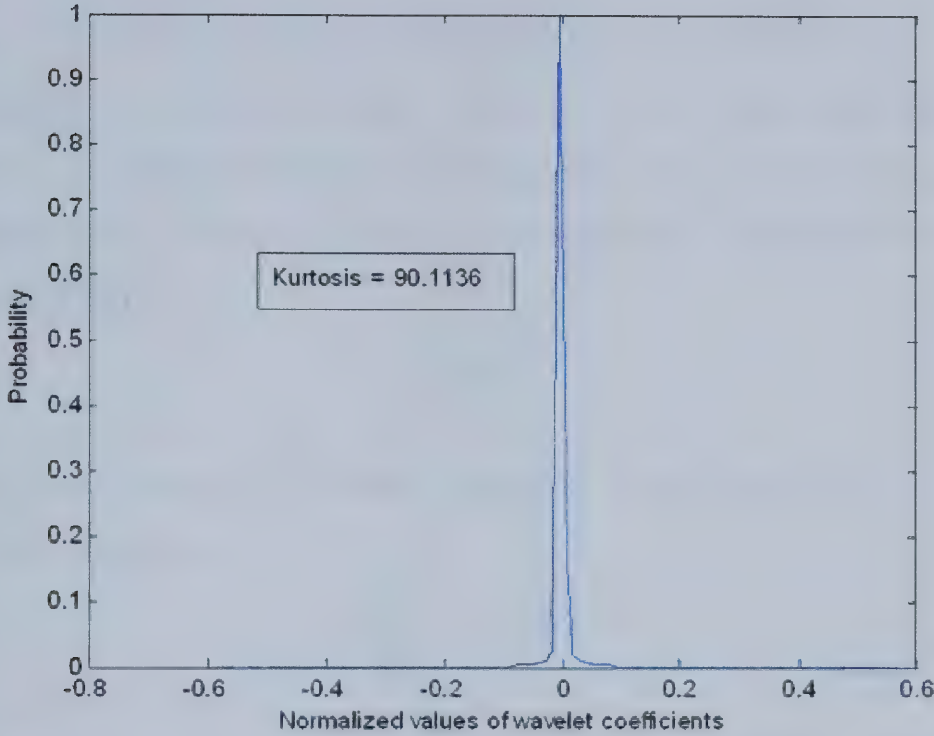


Figure 4.10 Probability distribution of the multi-scale shape representation of the shapes in the training set (Kurtosis = 90.1136)

Next, we use the training set of shapes to obtain a jointly-optimal multi-scale representation. We evaluate the significance of each basis function in representing the shapes in the training set. This process includes a feature-selection step in the wavelet-packet domain. The parsimonious representation of the shapes in the wavelet packet domain suggests that only a few coefficients are needed to represent a given shape. This sparsity also implies that the wavelet packet basis functions are not equally important for the statistical modeling task. Feature selection is achieved by discarding the basis functions that do not seem to contribute significantly to the overall energy of the population represented by the training set. The basis function whose coefficients tend to be constant is also less significant than the one whose coefficients exhibit more variations over the training set.

Let $X_i = [x_i(1) \dots x_i(N), y_i(1) \dots y_i(N)]^T$ be a zero-mean landmark-based representation of the i^{th} shape signal, with N vertices. x_i and y_i represent the x - and y -coordinates respectively.

Let $C_i = [C_i^x(1) \dots C_i^x(N), C_i^y(1) \dots C_i^y(N)]^T$ be the representation of the shape signals x_i and y_i in the wavelet packet domain and $i = 1 \dots K$, where K is the number of shapes in the training set. The wavelet packet representation C_i is obtained by projection of the shape signal X_i onto the best wavelet packet basis ϕ_{best} :

$$C_i = \phi_{best}^T X_i. \quad (4.13)$$

Let E_i be the energy of the i^{th} shape in the training set represented in the multi-scale wavelet packet domain:

$$E_i = \sum_{n=1}^N [C_i^x(n)^2 + C_i^y(n)^2]. \quad (4.14)$$

The total energy of the shapes in the training set can then be defined as follows:

$$E_{total} = \sum_{i=1}^K E_i = \sum_{i=1}^K \left\{ \sum_{n=1}^N [C_i^x(n)^2 + C_i^y(n)^2] \right\}. \quad (4.15)$$

Using equations 4.14 and 4.15, we define a measure for the contribution of each basis function to the overall energy. The contribution of the n^{th} basis function is defined as

$$S(n) = \sum_{i=1}^K \left\{ \sqrt{[C_i^x(n)^2 + C_i^y(n)^2]} \right\} / E_{total}. \quad (4.16)$$

Next, we rank the basis functions according to the contribution measure $S(n)$. We retain a sub-set of the basis functions that contribute to a percentage of 99.98 % of the total energy of the shapes in the training set. This feature selection step improves the statistical significance of the model. Table 4.1 shows the percentage of the energy captured for different target compression percentages when the size of the training set is fixed at 25 shapes. A compression percentage of 60 % is achieved with only 0.0526 % loss of the total energy. The mean square reconstruction error is kept below 0.5 pixels.

Table 4.1 Total energy captured for a number of target compression percentages, with the size of training set kept fixed to 25.

Target comp. percentage	35	40	45	50	55	60
% of energy captured	99.992%	99.9882%	99.9827%	99.975%	99.9640%	99.9474%

Figure 4.11 shows an example of a shape reconstructed by using the optimal wavelet-packet representation. This figure reveals that the sparsifying nature of the wavelet-packet basis leads to a useful dimension-reduction property without introducing significant reconstruction errors.

Within the context of anatomical shape modeling, the proposed multi-scale shape representation has the following merits:

- First, constructing a statistical model based upon the proposed representation will enhance the model generalization, with a limited number of training samples.
- Second, the feature selection step will also have a positive impact on the statistical significance and interpretability of the model.

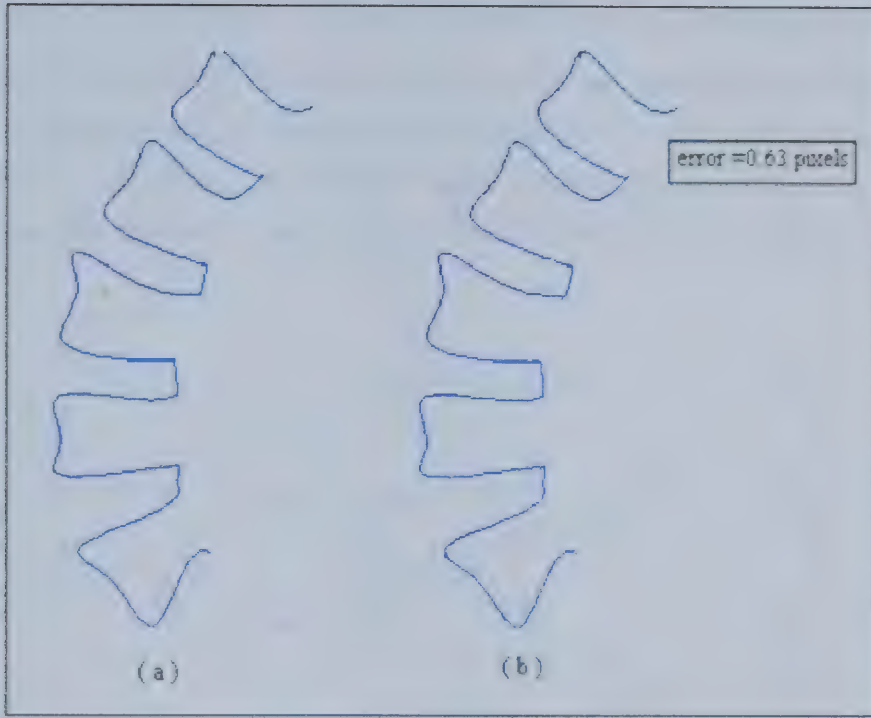


Figure 4.11 An example for shape reconstruction using the proposed multi-scale representation. (a) original shape, (b) reconstructed shape from the optimal multi-scale space (50% reduction).

4.5.2 Constructing the shape model

Using the multi-scale shape representation described in section 4.5.1, we next attempt to reveal the hidden independent sources of variations within the training set. First, we group the correlated wavelet packet coefficients into separate bands by using the spectral clustering technique [113]. The coefficients in the same band represent the areas of the shape that have correlated variations in the population. Next, we use Independent Component Analysis (ICA) to capture the localized variations in each band.

4.5.2.1 Spectral clustering

Prior to modelling the wavelet packet coefficients, we group the coefficients into a number of bands. The correlated coefficients are grouped together, while the coefficients across the bands tend to have minimum cross-correlation. This procedure is done by using the spectral clustering technique in the wavelet-packet domain [113].

The spectral clustering method tries to group highly correlated coefficients. Clustering is viewed as a graph-partitioning problem. The elements to be clustered are viewed as the nodes of a graph. Any two nodes $\{n_i, n_j\}$ are connected together by an undirected edge

with a weight w_{ij} . The weight on the edge reflects the degree of dissimilarity between the two nodes. Clustering is then achieved by finding the optimal partitioning for the similarity graph by using the Maximal Cut algorithm [113], which divides the graph into a number of sub-graphs, such that the nodes within a sub-graph have the highest covariance, and the nodes across the sub-graphs have the lowest covariance (see Figure 4.12).

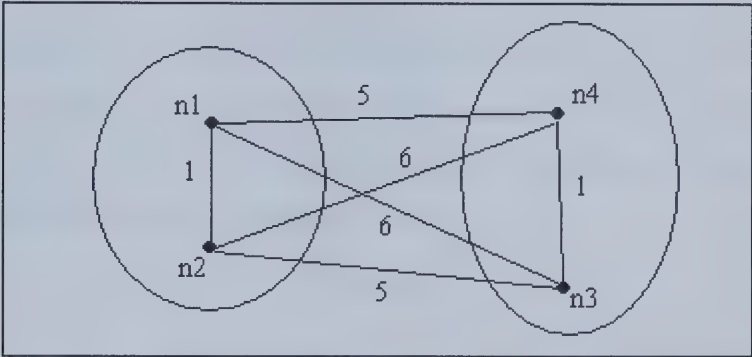


Figure 4.12 Illustration of the spectral clustering algorithm - undirected graph representation with nodes (n1, n2, n3, n4), and weighted edges

Given the training set of the shapes in the optimal wavelet-packet domain, we use the spectral clustering technique to cluster correlated wavelet packet coefficients. A fully connected undirected graph is first constructed with the wavelet packet coefficients as the nodes and the weights of the edges (w_{ij}) defined as the covariance among the coefficients in node (i, j) . By recursive graph partitioning of the coefficients, the highly correlated coefficients are grouped together into separate bands, with a minimum cross-covariance for the coefficients across the bands. The outcome is a total of B bands.

4.5.2.2 Statistical modeling using Independent Component Analysis (ICA)

After we have discovered the bands of the correlated coefficients in the wavelet-packet domain, the next step is to uncover hidden modes of variations in each band. This task is performed by modeling the coefficients in each band by using Independent Component Analysis (ICA).

Independent Component Analysis

Independent Component Analysis (ICA) is a generative modeling technique in which an observed multi-dimensional signal can be approximated as a weighted sum of the components that are as statistically independent as possible. ICA is often used to separate a group of mixed signals into a number of independent sources (independent components) with no knowledge about the mixing or the un-mixing matrix. Since its introduction, ICA has proven useful in a wide range of applications such as image feature extraction, denoising, blind source separation, and Functional MRI data processing [116]. Recently, ICA has been introduced as a useful tool in modeling anatomical structures [49].

In mathematical terms, the ICA modeling problem can be formulated as follows: Let X be a matrix with a number of observation signals coming from mixtures of unknown source signals (Independent Components -ICs):

$$X = AS , \quad (4.17)$$

where A is a matrix containing the mixing parameters, and S is the matrix of the source signals or independent components. The goal is to find the original source signals from the mixture X by estimating the de-mixing matrix U :

$$\hat{S} = UX . \quad (4.18)$$

Independent Component Analysis (ICA) seeks to extract statistical independent source signals by minimizing the statistical dependence among the components of the observed dataset (X). Several cost functions can be used to quantify the statistical dependence among the components of X . Two popular cost functions are the kurtosis and the mutual information [116]:

$$\text{Kurtosis} (X) = E(X^4) - 3 (E(X^2))^2 \quad (4.19)$$

$$I(x_1, x_2, \dots, x_n) = \sum_{i=1}^m H(x_i) - H(X) . \quad (4.20)$$

ICA-based modeling

In the context of statistical shape modeling, ICA tends to capture more localized shape variations than PCA, which is restricted to second-order statistics. The proposed multi-

scale statistical shape model uses ICA to model the wavelet packet coefficients in each sub-band.

Let C_i^l denote a vector of the wavelet-packet coefficients in the sub-band b_l for shape i .

For every sub-band, we calculate the mean coefficient vector over the training set:

$$\overline{C^l} = \frac{1}{K} \sum_{i=1}^K C_i^l, \quad (4.21)$$

where $\overline{C^l}$ is the average coefficient vector in sub-band l .

Next, we consider the zero-mean coefficients in each sub-band b_l as a mixture of the independent source signals with an unknown mixing matrix A . Any ICA algorithm can then be used to estimate the independent source signals by using the de-mixing matrix. We employ the FastICA algorithm to estimate the independent sources [117].

Let C^l be the matrix of wavelet packet coefficients in sub-band b_l for all shapes in the training set:

$$C^l = [C_1^l, \dots, C_K^l]. \quad (4.22)$$

Then the Independent Component Analysis (ICA) problem can be formulated as follows:

$$\begin{aligned} dC^l &= C^l - \overline{C^l} \\ dC^l &= A S^l \\ \hat{S}^l &= W dC^l \end{aligned}, \quad (4.23)$$

where S^l is the matrix with the independent components (ICs) in sub-band b_l , A is the mixing matrix, W is the de-mixing matrix, and $\hat{S}^l$ is the estimated matrix of ICs in sub-band b_l .

Figure 4.13 summarizes the procedure for constructing the multi-scale Wavelet-ICA basis of the proposed model. The complete multi-scale statistical shape model then includes the

joint best-basis matrix (B_{best}), the membership of each coefficient in a specific band, the average of the wavelet packet coefficients in each sub-band ($\bar{C}^{(l)}$), and the estimated independent source matrix for each sub-band ($\hat{S}^{(l)}$). The proposed multi-scale shape model can then be written as in equation 4.24, with $\hat{S}^{(l)}$ as the ICA basis in sub-band l :

$$\hat{C}^{(l)} \cong \bar{C}^{(l)} + \hat{S}^{(l)} b^{(l)}, l = 1..B. \quad (4.24)$$

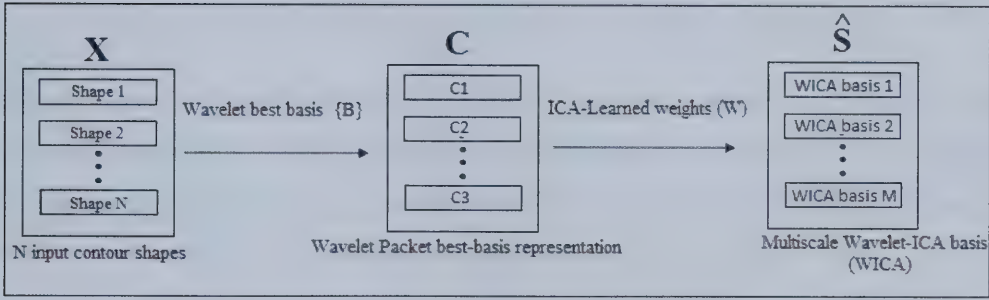


Figure 4.13 Construction of multi-scale independent Wavelet-ICA basis (WICA)

4.6 Experiments

In this section, we evaluate the performance of the proposed multi-scale shape model. We use a dataset of 40 outlined contours of five vertebrae from x-ray images of the lumbar spine [108]. Figure 4.14 shows an example of x-ray images with outlined vertebrae contours. First, the proposed statistical model is constructed by using a training set of shapes. Next, we evaluate the ability of the model to reconstruct unseen examples. Throughout this section, the new multi-scale shape model is abbreviated as BB-ICA. Experiments were conducted to achieve the following goals:

- Illustrate the ability of the new statistical model (BB-ICA) to capture local and global modes of deformations.
- Compare the performance of the proposed model against
 - i- The conventional PDM model (PCA-based) [35,36].
 - ii- Another related model from the literature that is based upon PCA analysis in the wavelet domain (BB-PCA) [80].
- Investigate the robustness of the new model to noise.

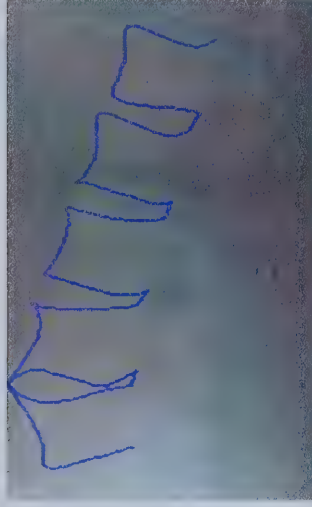


Figure 4.14 Example of x-ray images with outlined vertebrae contours [108]

We use Daubechies wavelets with six vanishing moments to build a six-level wavelet packet tree. The optimal shape representation for the shapes in the training set is obtained by using the method described in section 4.5.1. The FastICA algorithm is used to perform the ICA-based analysis in the clustered wavelet packet sub-bands [117].

4.6.1 Modes of deformations

This section evaluates the ability of the new model to capture local and global deformations.

Statistical independence

We use the mutual information as a measure of the statistical independence between the basis functions (or modes). The mutual information between two basis functions φ_1 and φ_2 is calculated as follows:

$$I(\varphi_1, \varphi_2) = H(\varphi_1) + H(\varphi_2) - H(\varphi_1, \varphi_2), \quad (4.25)$$

where $H(\varphi_i) = -E[\log(P_{\varphi_i}(\varphi_i))]$.

Figure 4.15 compares the statistical independence for the PCA basis, ICA basis, and the basis of the proposed model (BB-ICA). The PCA basis clearly has higher-order dependencies. In comparison to the ICA basis, the proposed model reduces the average mutual information by almost 60 %.

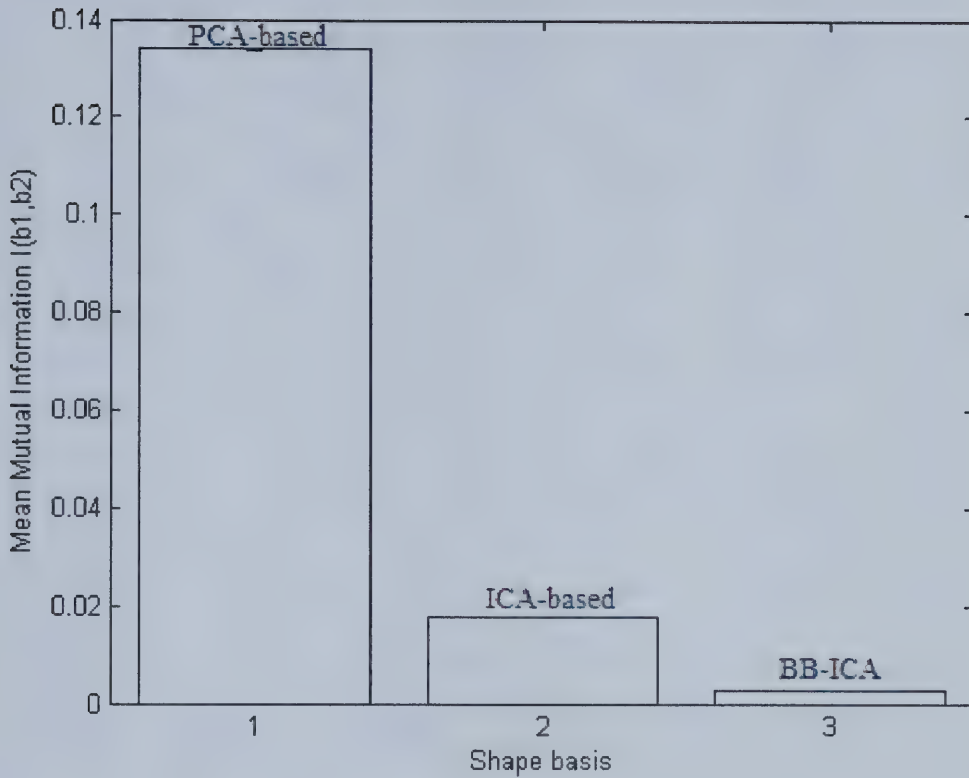


Figure 4.15 Evaluation of the statistical independence of the basis functions for PCA-based model, ICA-based, and the proposed BB-ICA model.

Multi-scale modes of deformations

Figure 4.16 shows the different modes of variations obtained by using the proposed multi-scale model. This figure reveals that the model is capable of capturing both the local and global deformations. The first mode (IC1) tends to capture the global variations in the lumbar spine. In the subsequent modes, we can observe more localized pathology-related deformations.

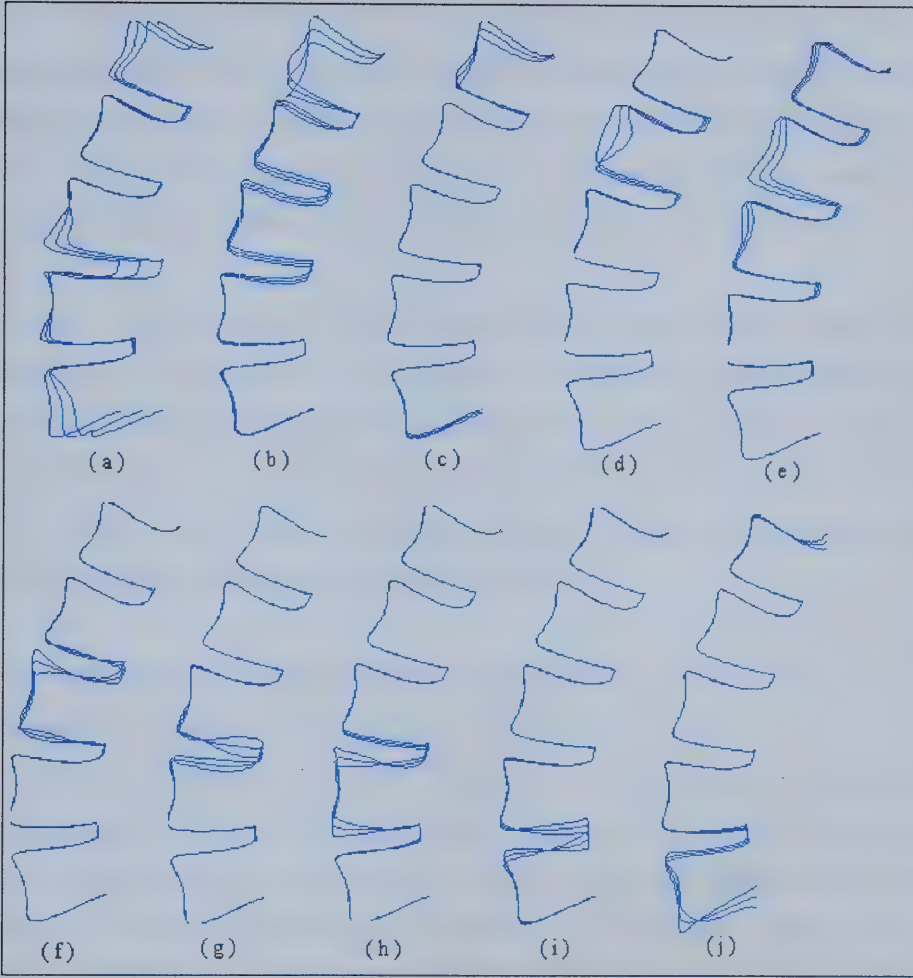


Figure 4.16 Multi-scale modes of the deformations of the proposed multiscale shape model.

4.6.2 Shape estimation

Next, we conduct experiments to evaluate the ability of the new model to reconstruct shapes not seen in the training set. We use the leave-one-out experiment to compare the performance of the new model (BB-ICA) against

- i- The conventional PCA shape model (PCA) [35].
- ii- A recently proposed model based upon PCA analysis of the wavelet sub-bands (BB-PCA) [80].

The leave-one-out experiment is conducted for several times by using a set of T training samples, where $T = [5, 10, 15, 20, 25, 30, 35, 40]$. The model is constructed by using any $T-1$ training samples, and we measure the reconstruction error for the left-out sample by using the constructed model. The reconstruction error between the estimated shape

Rec_i and the ground truth shape GT_i is defined as the root mean square error between the vertices of the ground truth and the reconstructed shape. For a given number of training samples, the overall reconstruction error of the model is defined as the average reconstruction error. This procedure is repeated for the three models under investigation: PCA [35], BB-PCA [80], and the proposed model (BB-ICA).

Figure 4.17 shows the plots for the average reconstruction error calculated for the different sizes of the training set. The curves in this figure reveal that the new model outperforms the other two approaches, even with few training samples available. The new model achieves an average error reduction of 30 % and 17 % over the PDM and the BB-PCA models, respectively. Thus, the proposed multiscale model appears to better generalize to the deformities unseen in the training set.

Figure 4.18 shows the reconstruction of a ground truth test shape by using the PCA model and the proposed multi-scale model (BB-ICA). Both models were constructed by using 15 training samples. While the PCA model seems to be unable to capture the local deformations in the vertebrae bodies, our model appears to more closely follow the test shape. Figure 4.19 shows a zoom-in to the reconstruction of two deformed vertebrae by using both the new multi-scale model and the PCA-based model. Again, the proposed shape model tends to more closely follow the original contour.

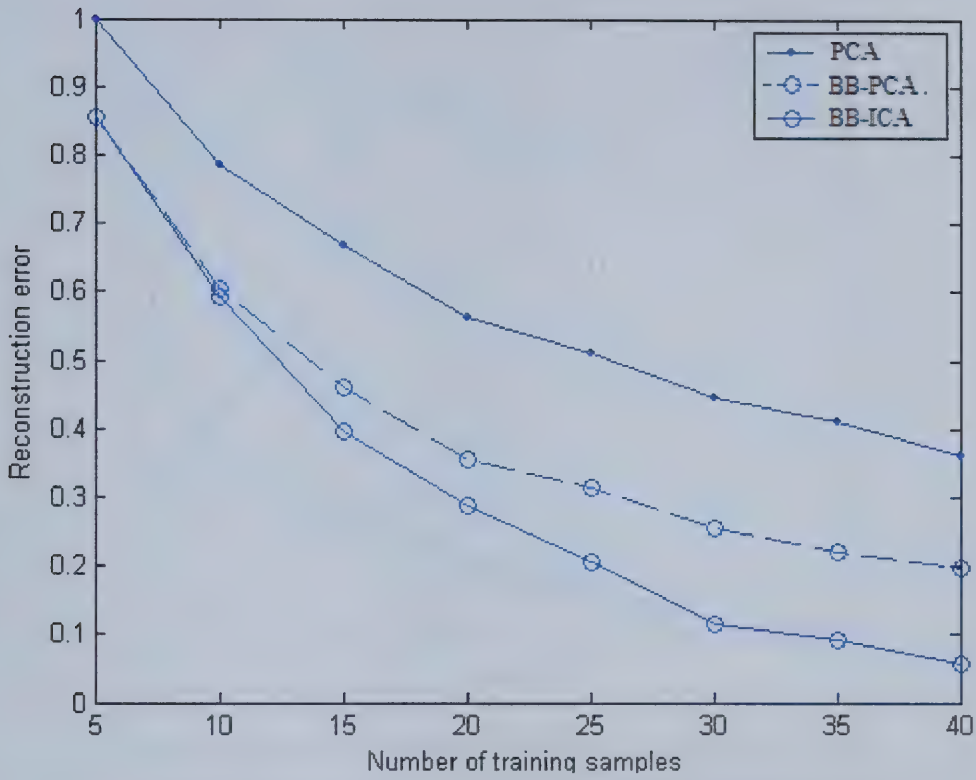


Figure 4.17 Average reconstruction error for the following shape models: PCA-based, best-basis PCA (BB-PCA), and the proposed multi-scale best-basis ICA model (BB-ICA)

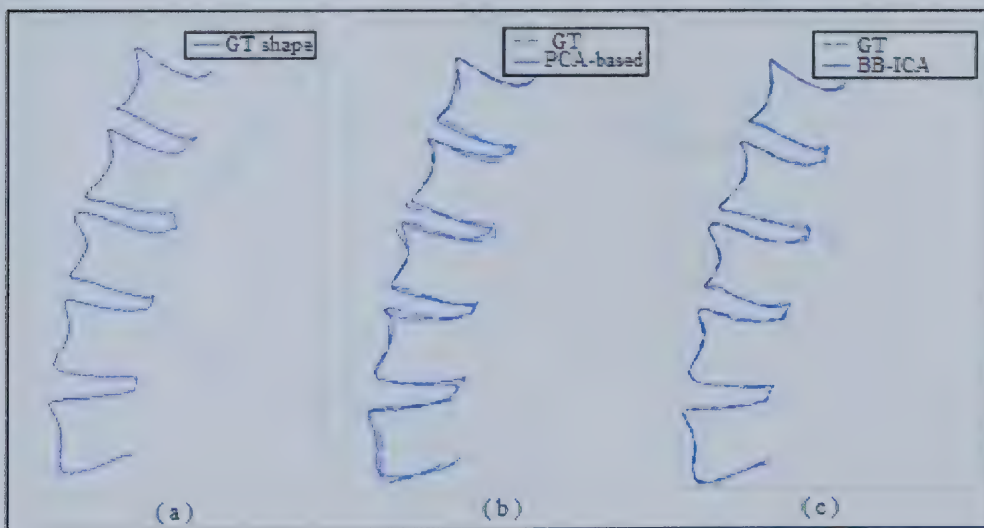
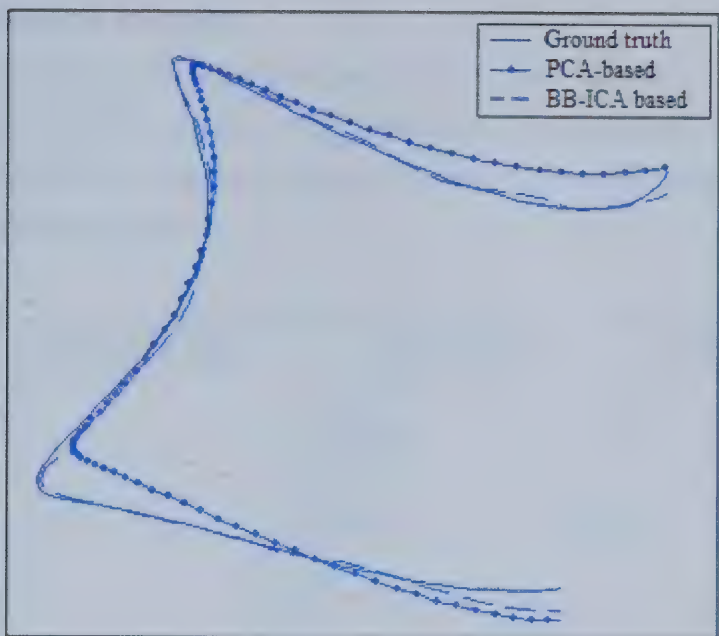
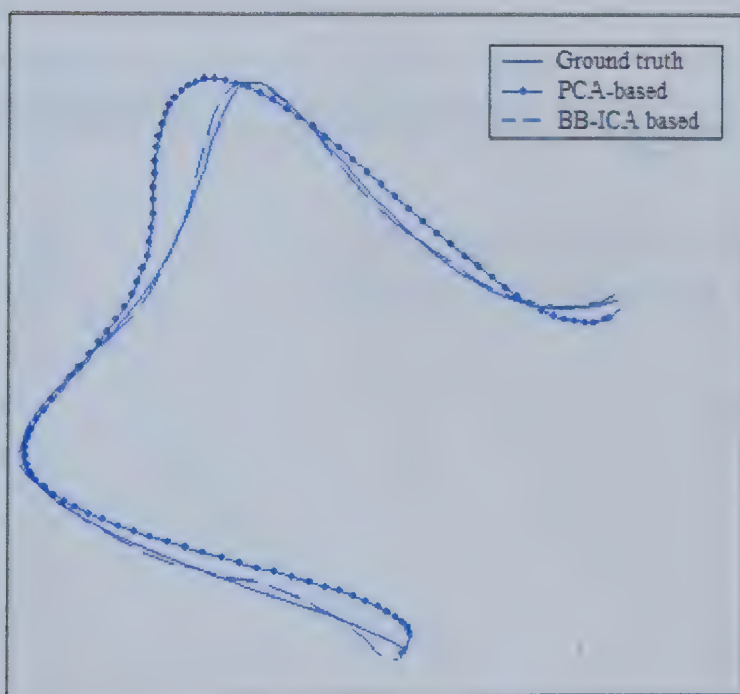


Figure 4.18 Reconstruction of a ground truth test shape with models constructed by using 15 training samples: (a) GT, (b) PCA-based model, (c) BB-ICA model



(a)



(b)

Figure 4.19 Zoom-in to the reconstruction of two L2 lumbar vertebrae by using BB-ICA, and PCA-based models

4.6.3 Robustness to noise

Finally, we evaluate the robustness of the proposed shape model to noise by adding Gaussian noise to the test shapes and calculating the average reconstruction error by using the leave-one-out experiment. Figures 4.21 and 4.22 show that the proposed model is minimally affected by noise.

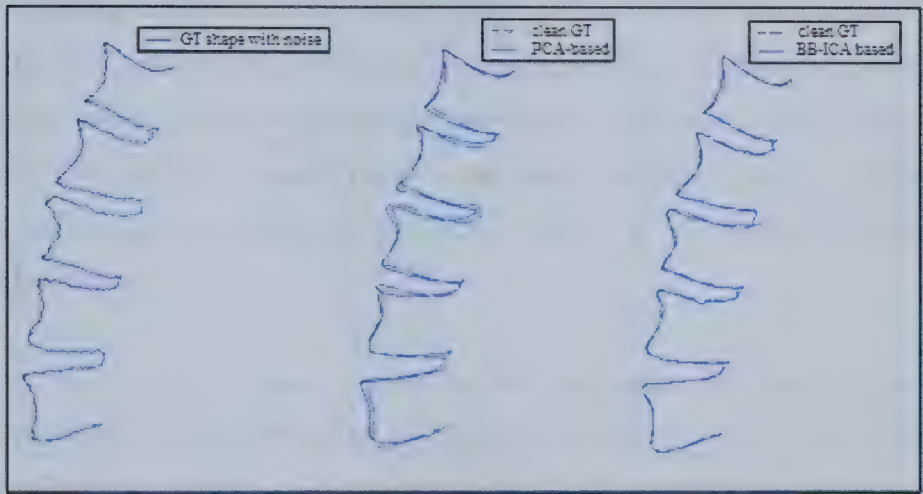


Figure 4.21 Reconstruction from a noisy test shape by using 20 training samples. (a) Ground truth noisy shape, (b) PCA-based reconstruction, (c) BB-ICA based reconstruction.

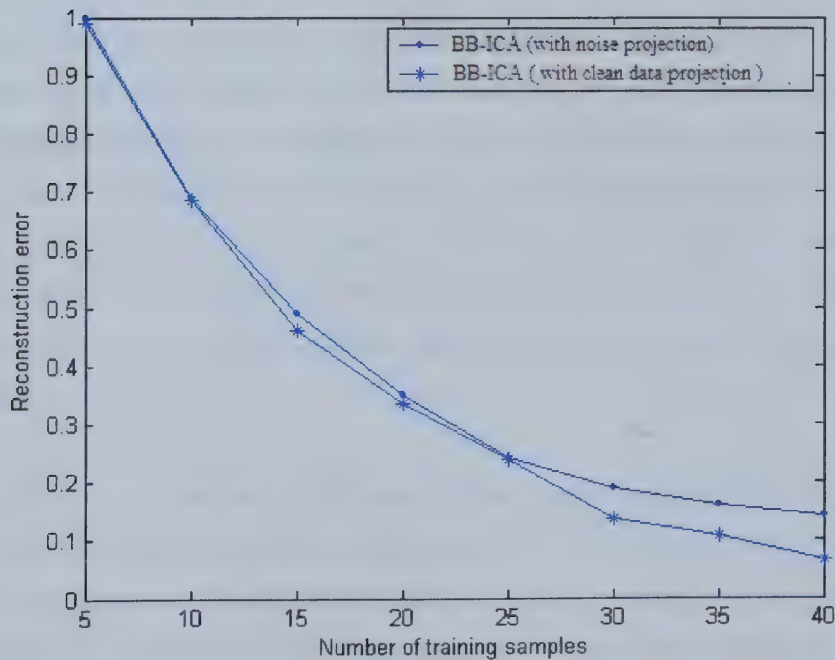


Figure 4.22 Effect of noise on the reconstruction error of the proposed model.

4.7 Discussion

In this chapter, we presented a novel multi-scale shape model for modeling the local and global variations in anatomical structures. Our model integrates the concepts of multi-scale decomposition, sparsity, feature selection, and statistical independence to extract sets of multi-scale modes of deformations. This section discusses the main aspects of the new model in comparison to related work in the literature.

Compared to the conventional PCA-based Active Shape Model (ASM) [35], the proposed model has better generalization ability, particularly when few training samples are available. For a set of K training samples, our model can have up to $K - 1$ degrees of freedom from each sub-band, in comparison to a total of $K - 1$ degrees of freedom for the ASM model.

We also compare the proposed model and the approaches proposed by Nain [47], Davatzikos et al. [45], and Mahmoud et al. [80]. Our model uses ICA for wavelet packet sub-band analysis, whereas all three models in [47, 45, and 80] rely on PCA as the statistical modeling tool. By using PCA to analyze wavelet sub-bands, we are restricting the model to second-order statistics. However, wavelet coefficients tend to be sparse, with non-Gaussian distribution, particularly for higher frequency bands. This feature encourages using Independent Component Analysis (ICA) for the analysis of the wavelet packet sub-bands. ICA assumes that the coefficients in each sub-band are mixture observation signals made up of independent components. Using the central limit theorem, we can assure that since the observations (wavelet packet coefficients) are non-Gaussian, the sources will be even more non-Gaussian. This result implies that ICA would be more successful than PCA in searching for hidden independent modes of variations in the wavelet packet sub-spaces. This argument agrees with the experimental results reported in section 4.6.

In [47], Nain et al. introduced a multi-scale shape model using spherical wavelets and PCA. This model is limited to shapes with spherical topology, due to the use of spherical wavelet basis functions. On the other hand, our model constructs a multi-scale shape representation by searching for the best basis in an overcomplete dictionary of basis functions. This feature allows for more flexibility in modeling variations of the different shape topologies.

In [54], the authors presented a localized shape model using Sparse Principal Component Analysis (SPCA). In this model, locality is achieved by explicitly imposing sparsity constraint onto the loadings of a regular PCA. Doing so implies that the resulting shape model will still suffer from some of the inherent limitations of PCA-based models. On the other side, our model is capable of capturing both global and local variations by first projecting the data onto an optimal multi-scale sparse domain and then applying a higher-order statistical analysis technique.

4.8 Conclusion

This chapter tackled some of the challenges in the statistical modeling of anatomical structures. Throughout the chapter, we had two main goals:

- First, finding the proper shape representation to accurately characterize anatomical structures.
- Second, constructing an efficient statistical model to capture both the local and global pathology-related modes of variations.

We presented a novel multi-scale statistical shape model that uses the concepts of sparsity, dimension reduction, and statistical independence to extract sets of localized clinically- relevant modes of deformations. Within a joint best-basis selection framework, the proposed model benefits from the multi-scale and sparsifying nature of wavelet packets and the ability of Independent Component Analysis (ICA) to capture higher-order statistics. The new method was applied to model variations in the human lumbar spine. Our experiments have demonstrated the ability of the proposed model to capture both local and global deformations. The novel multi-scale shape model can be beneficial in a wide range of medical imaging tasks. In the subsequent chapters, we utilize this shape model for the segmentation and quantification of vertebral structures in digital x-rays.

5 Vertebral Shape Analysis and Classification using Localized, Pathology-related Shape Parameters

5.1 Introduction

Statistical shape analysis is used to understand how anatomical structures tend to vary for various reasons such as the development of a certain pathology, aging, and surgery. A long-standing goal for medical imaging research has been to enhance the ability to distinguish between the anatomy of healthy individuals and those with pathology. This ability can allow for the accurate early detection and diagnosis of diseases. It also opens the door for a wide range of applications such as surgical planning and digital archiving. Despite the research efforts in the field, the efficient analysis and classification of anatomical structures continues to be a formidable task. Anatomical structures often exhibit a wide range of variations due mainly to inter-subject variability and structural differences among different pathologies.

This chapter focuses on the analysis and classification of human vertebral structures. Radiographs of the human spine are frequently examined to assess vertebral abnormalities. Medical experts often examine hundreds of spine x-rays to determine the existence of various vertebral pathologies. The common pathologies of interest are anterior osteophyte, disc space narrowing, and fractures. By carefully inspecting the outline shapes of the vertebral bodies, experts are able to identify and assess vertebral abnormalities with respect to the pathology under investigation. Features like osteophyte (bony growths on vertebral corners) and disc space narrowing are often used as visual evidence of osteoarthritis (degenerative joint disease). Morphometric measurements of vertebral bodies are often used as evidence of vertebral fractures and osteoporosis (a skeletal disease characterized by a decline in bone mass). Traditionally, vertebral shape analysis has been performed by using a few points that outline vertebral bodies. There is now a growing trend towards shape-based methods for better quantification of vertebral deformities. Most vertebral shape analysis and classification schemes rely on global error-prone representations. However, most pathologies tend to be localized. Thus, new shape analysis and classification methods are needed to capture subtle pathology-related information.

In this chapter, we present a novel framework for the analysis and classification of vertebral structures using pathology-related Wavelet-ICA shape parameters. For vertebral shape analysis, we use the parameters of the multi-scale shape model proposed in chapter four, in combination with regression analysis techniques. Using the proposed framework, we are able to establish a statistically relevant relationship between the Wavelet-ICA modes of deformations and clinical information. This ability allows for the anatomically meaningful interpretation of the multi-scale modes of variations. Furthermore, we present a novel method for computer-assisted multi-vertebrae classification. The new classification scheme incorporates the following steps: localized shape modeling, feature selection, and statistical classification. We demonstrate the ability of our proposed method to identify local shape changes in vertebral structures related to various pathologies. Thanks to the locality of the features, the new classification scheme can assess the condition of several vertebrae simultaneously.

The rest of the chapter is organized as follows: Section 5.2 presents the details of the proposed shape analysis and classification framework. Our experiments and results are presented in section 5.3. Discussion and conclusions are presented in sections 5.4, and 5.5, respectively.

5.2 Methodology

5.2.1 Statistical analysis of vertebral structures by using Wavelet-ICA representation.

We present a novel framework for the analysis of vertebral structures by using the parameters of the multi-scale shape model proposed in chapter four. We refer to the shape parameters as the Wavelet-ICA representation (WICA). The proposed framework incorporates the following: rich multi-scale representation of vertebral shapes, feature selection approaches, and statistical analysis techniques. We aim to achieve the following goals:

- a. Discriminate between two groups of normal and pathological vertebrae.
- b. Establish a relationship between the localized modes of variations and the clinical outcome variables.

Figure 5.1 shows a block diagram of the proposed vertebral analysis framework using Wavelet-ICA representation (WICA). We also compare the proposed framework against a PCA-based point distribution shape analysis scheme.

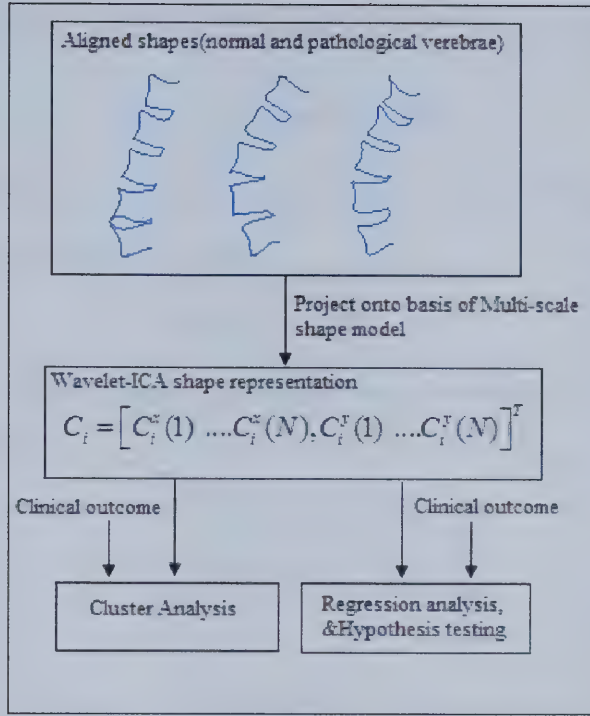


Figure 5.1 Vertebral Shape Analysis framework by using Wavelet-ICA representation

5.2.1.1 Establishing clinical relevance by using regression analysis

We use regression analysis to establish the relationship between the Wavelet-ICA shape parameters (WICA) and a clinical outcome variable.

Regression analysis is a technique that is normally used to assess the relationships between two or more variables. The relationship between the variables is generally expressed as

$$y = f(x) + \mathcal{E}, \quad (5.1)$$

where x is called the independent variable (predictor), y is the dependent variable (response), and $\mathcal{E}$ is a random error.

The WICA-based regression analysis framework

We use a regression model to establish the relationship between the WICA shape parameters and a clinical dichotomous outcome variable which can have two values: either 0 or 1. “0”s correspond to a normal vertebra, and “1”s correspond to a vertebra with a pathology.

Let Φ_{WICA} be the matrix of the multi-scale Wavelet-ICA localized basis functions:

$$\Phi_{\text{WICA}} = \begin{bmatrix} \phi_1^T & \phi_2^T & \dots & \phi_M^T \end{bmatrix}. \quad (5.2)$$

The projection of the shape x_i onto the set of basis functions Φ_{WICA} is referred to as the Wavelet-ICA representation (WICA):

$$\mathbf{y}_i = [y_{i1} \ y_{i2} \dots y_{iM}]^T. \quad (5.3)$$

For a number of N training samples, we construct a score matrix $\mathbf{Y}$ of the Wavelet-ICA shape representation:

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1M} \\ y_{21} & y_{22} & \dots & y_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ y_{K1} & y_{K2} & \dots & y_{KM} \end{bmatrix}, \quad (5.4)$$

where y_{ij} is the projection of the shape x_i onto the Wavelet-ICA basis ϕ_j^T . K is the size of the training set and M is the number of wavelet-ICA basis.

Let C_i be the clinical outcome variable for the i^{th} vertebra in the training set. We apply a series of univariate regression tests to establish the relationship between each deformation mode, represented by the corresponding shape parameter (y_{ij}), and the clinical outcome variable C_i :

$$\mathbf{y}_j = \beta_j \mathbf{c} + \varepsilon, \quad (5.5)$$

where $\mathbf{y}_j$ is the continuous dependent variable which represents the projections of all shapes in the training set onto the j^{th} basis. $\mathbf{c} = [c_1, c_2, \dots, c_K]$ is the binary independent variable that represents the clinical information. β_j is the regression coefficient which encodes the mean difference between the groups of the normal and pathological vertebrae on the continuous outcome variable.

We use the least square method to find the regression coefficients, which are the ones that minimize the sum of the squared residual error (RSE):

$$RSE(\beta_j) = \sum_{i=1}^K (y_{ij} - \beta_j c_i)^2 \quad (5.6)$$

$$\hat{\beta}_j = \arg \min_{\beta_j} \{RSE(\beta_j)\}. \quad (5.7)$$

The minimization problem is solved by differentiating equation 5.6 with respect to β_j , and setting the results equal to zero:

$$\hat{\beta}_j = \frac{\sum_{i=1}^K y_{ij} c_i}{\sum_{i=1}^K c_i^2}. \quad (5.8)$$

Hypothesis testing

Next, we perform hypothesis testing by using the non-parametric method [118]. This testing procedure evaluates the statistical significance of the regression coefficients. The result of this test is a range of t-scores with the corresponding p-values which is the probability that a significant relationship is declared when the variables are unrelated. Usually, a p-value of less than 0.05 indicates a significant relationship between the model variable (i.e., its deformation modes) and the clinical outcome variable. The non-parametric permutation test is a type of hypothesis testing that does not need parametric reference distributions. In non-parametric testing, a reference distribution is obtained by calculating all possible values of the test statistic under the re-arrangement of the labels on the observed points. We also perform a 5% false discovery estimation procedure in order to control the expected proportion of false positives due to multiple comparisons. More details about non-parametric hypothesis testing methods can be found in [118,123].

Using regression coefficients and their corresponding p-values, we are able to establish a relationship between the localized modes of deformations of the multi-scale shape model and clinical information. This ability also reveals the clinical relevance and interpretability of the model proposed in chapter four.

5.2.1.2 Discriminant Analysis

We use class-discriminability to assess the ability of the Wavelet-ICA shape parameters to distinguish between normal and pathological deformations. For each deformation mode, we calculate the discriminability ratio (r) defined as the ratio between between-class variability ($\sigma_{between}$) and within-class variability (σ_{within}):

$$\sigma_{between} = \sum_i (\mu_i - \mu)^2 \quad (5.9)$$

$$\sigma_{within} = \sum_i \sum_j (y_{ij} - \mu_i)^2 \quad (5.10)$$

$$r = \frac{\sigma_{between}}{\sigma_{within}}, \quad (5.11)$$

where, μ is the overall mean of the coefficients over the training set. μ_i is the average of the coefficients for class i . y_{ij} is the coefficient of the j^{th} training shape in class i .

By calculating the discriminant ratio (r) for each deformation mode, we can evaluate the ability of the representation to distinguish between the classes of normal and pathological vertebra.

5.2.2 Multi-vertebrae classification

An essential step in a computer-aided diagnosis system is the automatic classification of a patient's pathology. Computer-assisted classification methods aim at reducing the time and cost of the error-prone and time-consuming manual classification process. In medical imaging, an efficient classification scheme includes the following:

- A rich representation that can capture important anatomical information.
- An efficient classifier that can distinguish between different classes.

The lower the dimension of the feature space, the more manageable is the classification stage. However, most pathologies are related to local deformations. Thus, the challenging task is to find a suitable low-dimensional representation that can still capture subtle pathology-related information. In this section, we propose a new multi-vertebral classification scheme using localized pathology-related shape parameters.

At the core of the proposed framework is a multi-scale localized shape representation based upon wavelet packets and Independent Component Analysis (ICA). The multi-scale shape model presented in chapter four can extract sets of localized modes of deformations specific to the vertebra under investigation. By projecting the shapes onto any specific set of deformation bases, we obtain low-dimensional features linked to the vertebra of interest. These features are then used as input to a Fisher classifier to classify the target vertebra as normal or abnormal.

This section outlines the details for the proposed classification scheme. A block diagram of the model is shown in Figure 5.2. The proposed method incorporates the following key features:

1. A localized multi-scale shape model that uses the concepts of sparsity, dimension reduction, and statistical independence to extract localized modes of deformations.
2. A set of pathology-related Wavelet-ICA shape parameters obtained via basis sorting and feature selection.
3. Simultaneous supervised classification for all the vertebrae of interest as normal or abnormal (i.e., with pathology).

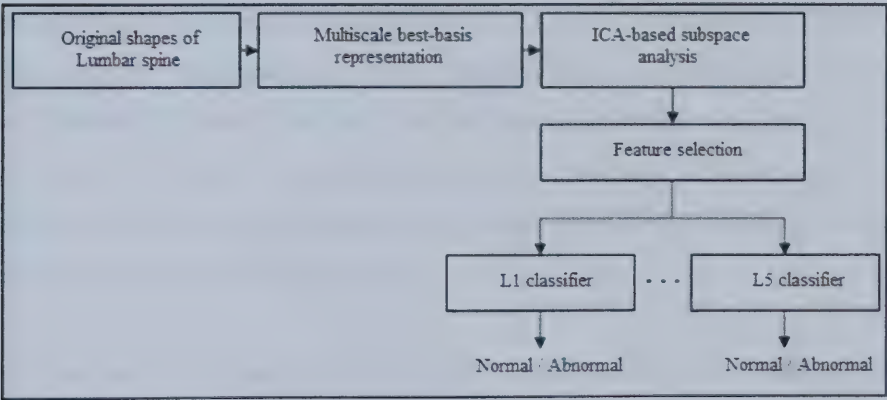


Figure 5.2 Block diagram for the proposed vertebral classification scheme.

The following summarizes the steps for the proposed method:

5.2.2.1 Multi-scale Wavelet-ICA shape representation

Let a zero-mean shape vector be represented as

$$X_i = [x_{i1} \dots x_{iN}, y_{i1} \dots y_{iN}]^T, \quad (5.12)$$

where x and y are x -and y -coordinates.

We construct the multi-scale shape model described in chapter four. Let Φ_{WICA} be the matrix of the multi-scale basis functions. In the shape space, Φ_{WICA} constitute sets of localized pathology-related deformations. The Wavelet-ICA representation $\mathbf{y}_i$ is the vector of coefficients obtained by projecting the shape X_i onto the Wavelet-ICA basis matrix Φ_{WICA} .

$$\mathbf{y}_i = [y_{i1} \ y_{i2} \dots y_{iM}]^T. \quad (5.13)$$

5.2.2.2 Feature selection

The main goal of this step is to find the subsets in the Wavelet-ICA representation that best capture the pathological information specific to every vertebra of the lumbar spine. The feature-selection step benefits from the sparsity and locality of the multi-scale model to obtain subsets of features that are not only small in dimension, but also spatially localized and have high discriminating power. This step aims at increasing the generalization ability of the classifier and also allows for efficient multi-vertebrae classification in the Wavelet-ICA domain.

First, we perform basis sorting according to spatial locality by arranging the basis functions into sets of localized deformations such that each set corresponds to a specific vertebra. Thus, by projecting the shapes onto any specific set, we obtain a low-dimensional feature vector that is linked to the vertebra of interest. We also use the outcome of the discriminant analysis performed in section 5.2.1.2, to further select the features with the highest discriminability according to equation 5.14. The outcome sets of feature vectors are then used as inputs to the classifier.

A- Feature Spatial Locality (FSL):

We arrange the basis functions into sets of localized deformations such that each set corresponds to a specific vertebra. The Feature Spatial Locality criterion is defined as follows:

Let $\hat{\mathbf{X}}_i$ describe the spatial shape variation vector corresponding to the i^{th} component of the WICA feature vector. Every element in the WICA feature vector corresponds to a particular WICA basis function or variation mode:

$$\hat{\mathbf{X}}_i = (\Phi_{BB})^{-1} [\mathbf{W}^{-1} * \mathbf{y}_i * \mathbf{e}_i], \quad (5.14)$$

where Φ_{BB} : is the matrix of the Wavelet Packet best basis set.

$\mathbf{W}$: is the learnt ICA weight matrix.

$\mathbf{y}_i$: is the WICA-feature vector of the input shape.

$\mathbf{e}_i$: is a vector with “1” at the i^{th} position and zero every where else.

$\mathbf{e}_i \in R^M$, M is dimension of the original shape space in spatial domain.

We partition the shape variation vector $\hat{\mathbf{X}}_i$ into five regions corresponding to the five lumbar vertebrae. In each region (l), we calculate the Euclidean distance between the

shape variation vector $\hat{\mathbf{X}}_i^{(l)}$ and the mean shape $\bar{\mathbf{x}}^{(l)}$:

$$d_i^{(l)} = \left\| \hat{\mathbf{X}}_i^{(l)} - \bar{\mathbf{x}}^{(l)} \right\|_2, l=1,2,3,4,5. \quad (5.15)$$

Given a Wavelet-ICA shape representation vector ($\mathbf{y}_i$), the i^{th} element of the feature vector is then labelled as being spatially localized to the lumbar vertebra L_k . k is given by equation 5.16:

$$k = \arg \max_l \{d_i^{(l)}, l=1..5\}. \quad (5.16)$$

Using this criterion, we extract sub-sets of the Wavelet-ICA features that correspond to each vertebra. Figure 5.3 shows an example of a WICA feature element spatially localized to the L4 vertebra.

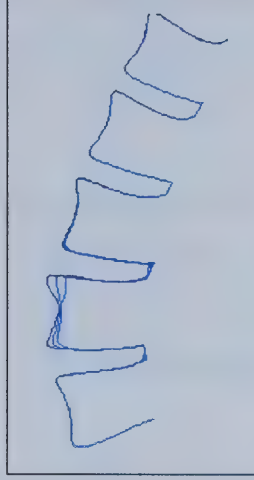


Figure 5.3 Example of a WICA-shape parameter spatially localized to L4 lumbar vertebra.

B- Feature Discriminatory Power (FDP):

For every vertebra, we use the discriminability metric presented in section 5.2.1.2 to select a subset of the Wavelet-ICA feature space that has the most discriminatory power.

C- Fusion of feature selection methods:

We also investigate the effect of fusion of the two previous criteria. First, we select the subset of features spatially localized to the vertebra under investigation. Second, we use the statistical discriminatory power criterion to further reduce the dimensionality of the feature space.

5.2.2.3 Shape-based classification

Next, we use the feature vectors obtained from the previous section, along with a suitable statistical classifier to distinguish between normal and pathological populations of the lumbar vertebrae.

With each vertebra being labelled as normal or abnormal, the input to the classification problem is a training set of pairs of feature vectors and the corresponding labels:

$$\{ (y_1, l_1) , (y_2, l_2) , ... , (y_O, l_O) \} , \quad (5.17)$$

where y_i is the feature vector for the i^{th} shape. l_i is a label that can either be “1” or “-1”, indicating whether the vertebra of interest is normal or abnormal (i.e., with a pathology).

A statistical classifier aims at dividing the feature space into regions separated by decision boundaries. A training set is used to learn the parameters of the classification model. Various classification methods can be found in the statistical learning literature [119, 124].

Generally speaking, there are two common approaches for classification.

The first approach, referred to as the elementary decision theory, uses the available training set to estimate the underlying class conditional probability functions, ($p(\mathbf{y} / c(\mathbf{y}) = 1)$ and $p(\mathbf{y} / c(\mathbf{y}) = -1)$). The classification is then achieved by using the Posterior Likelihood estimate:

$$c(\mathbf{y}) = \arg \max_i p(\mathbf{y} / c(\mathbf{y}) = i) p(c(\mathbf{y}) = i), \quad (5.18)$$

where $p(c(\mathbf{y}) = i)$ is the prior knowledge, $i=1$ for the normal case, and $i=-1$ for the pathological case.

The other classification approach involves using decision functions to estimate decision boundaries. Instead of imposing assumptions about the class posterior probability, this approach makes assumptions about the discriminant function. A discriminant function is defined as follows:

$$f(\mathbf{y}) + T \begin{cases} >0 \\ <0 \end{cases} \Rightarrow \mathbf{y} \in \begin{cases} c1 \\ c2 \end{cases}, \quad (5.19)$$

Where T is the discrimination threshold.

In the proposed vertebrae classification scheme, we use the Fisher classifier to classify a specific vertebra as normal or with osteophyte pathology.

Fisher Linear Discriminant Classifier:

The Fisher Linear discriminant classifier tries to separate the two classes of normal and pathological vertebrae by finding the direction ($\mathbf{w}$) along which the two classes are best separated according to the Fisher criterion, which is defined as the ratio of the between-class to within-class variances:

$$J_{Fisher} = \frac{|\mathbf{w}^T (\mu_1 - \mu_2)|^2}{\mathbf{w}^T S_w \mathbf{w}}, \quad (5.20)$$

$$S_w = \frac{1}{n-1} \left[n_1 \sum_1 - n_2 \sum_2 \right], \quad (5.21)$$

where $\mathbf{W}$ is the direction of maximum discrimination,

μ_1, μ_2 are the mean values for the class of normal and pathological vertebrae respectively.

S_w is the within-class scatter matrix.

$\sum_1$ and $\sum_2$ are the covariance estimates for the normal and pathological classes.

We close this section by commenting that the common challenge among various classification methods is the high dimensionality of the feature space with respect to the size of the training set available. In this chapter, we try to tackle this issue through the locality of the shape model, and the feature selection stage. Sets of low-dimensional feature vectors are used to describe each vertebra of the lumbar spine.

5.3 Experiments and Results

In this section, we evaluate the performance of the proposed vertebral shape analysis and classification schemes. Experiments are conducted by using a dataset of lateral x-ray images of the human lumbar spine [108]. For each x-ray image, the lumbar vertebrae L1-L5 are annotated by experienced radiologists. Each vertebra is labelled as normal or abnormal (i.e., with osteophyte pathology). We focus on the anterior osteophyte (AO) as the pathology of interest. Anterior osteophytes are bony growths along the anterior edge of the vertebra bodies. Figure 5.4 shows an example x-ray image of a lumbar spine with anterior osteophyte pathology.

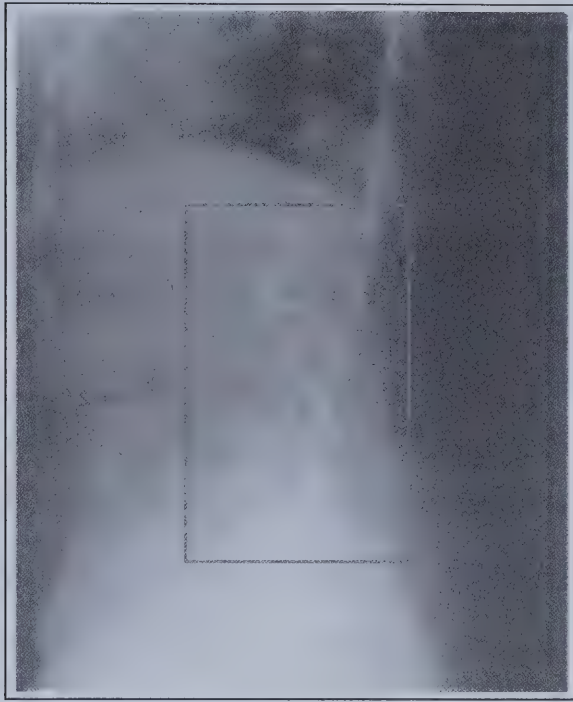


Figure 5.4 Lumbar spine x-ray image from the National Library of medicine dataset [108].

We use a set of 150 outlined contours of lumbar vertebrae L1-L5, with each vertebra labelled as normal or abnormal. Figure 5.5 shows vertebrae with anterior osteophytes. 50 % of the vertebrae were labelled as normal, and the rest as abnormal. A set of 100 contours is used for training, and the remaining 50 contours are used for testing. The following summarizes the steps of the classification scheme:

1. Construct the multi-scale shape model, and obtain the shape parameters, as described in section 4.3.
2. Use the shape parameters of the model as feature vectors. We refer to these shape parameters as Wavelet-ICA (WICA) features.
3. Use the training set to calculate the mean and standard deviation of the feature vectors for the two classes of vertebrae (normal and pathological).
4. Normalize the feature vectors in the training set.
5. Use the set of feature vectors to train the Fisher classification rule.
6. Evaluate the accuracy of the classification scheme.
7. Repeat steps 1-5 for 20 randomly generated training and testing sets.

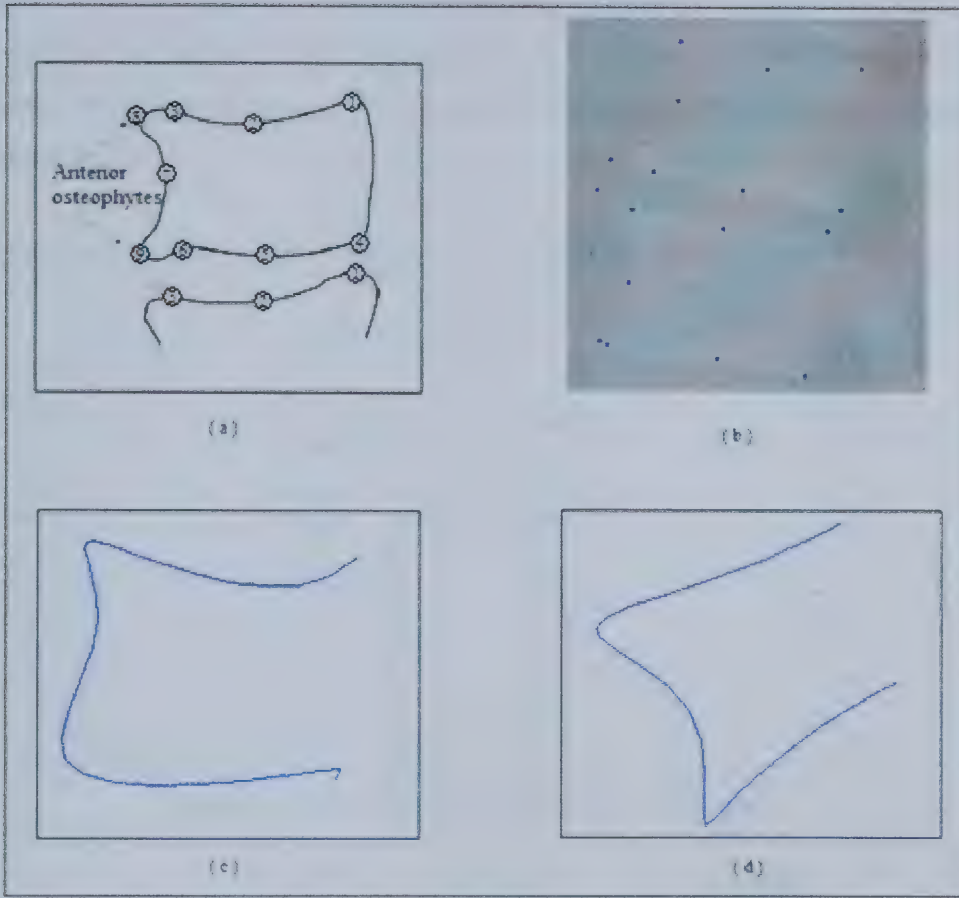


Figure 5.5 Vertebrae with anterior osteophytes. (a) Label points selected by expert radiologist. Points 8, 9 characterize osteophyte pathology, (b) two lumbar vertebrae with the 9-points model (c) Contour of Lumbar vertebra with upper left-hand corner osteophytes, (d) Contour lumbar vertebra with upper and lower left-hand corner osteophytes.

The experiments are conducted to achieve the following goals:

- Evaluate the ability of the multi-scale shape model to predict pathological vertebral deformities.
- Evaluate the ability of the proposed Wavelet-ICA shape analysis scheme to explain the anatomical differences between normal and pathological vertebral populations.
- Evaluate the accuracy of the proposed multi-vertebrae classification scheme.

5.3.1 Shape prediction in presence of pathology

We investigate the ability of the multi-scale shape model to capture localized subtle pathology-related variations. Figure 5.6 shows examples of deformation modes that capture variations in the L4 vertebra of the lumbar spine. Using the leave-one-out experiment, we investigate the ability of the model to predict unseen vertebrae with

various degrees of anterior osteophytes pathology. Figure 5.7 compares the prediction capability of the proposed model, and that of a PCA-based model [36]. Our model tends to better reconstruct the shapes of normal and abnormal vertebrae (i.e., with anterior pathology).

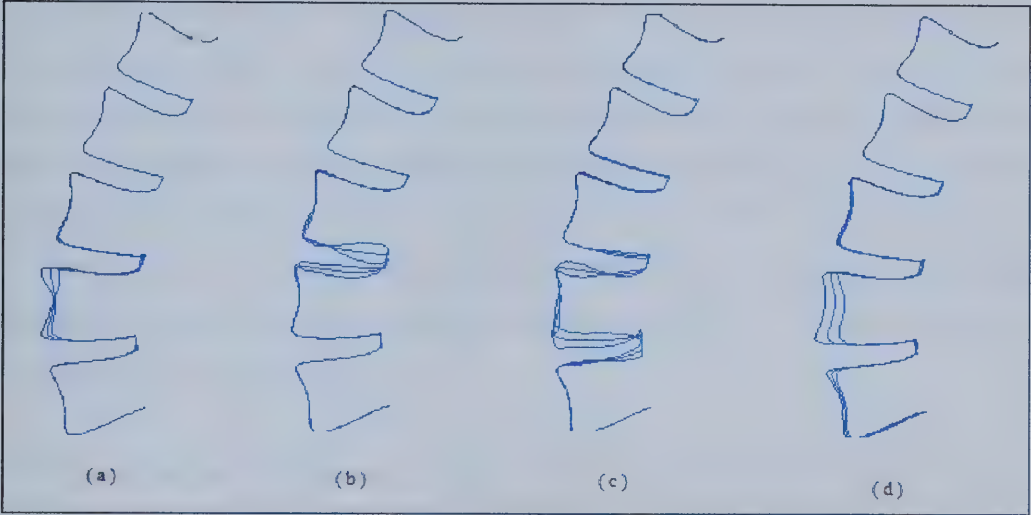


Figure 5.6 Local modes of the deformations for L4 vertebra

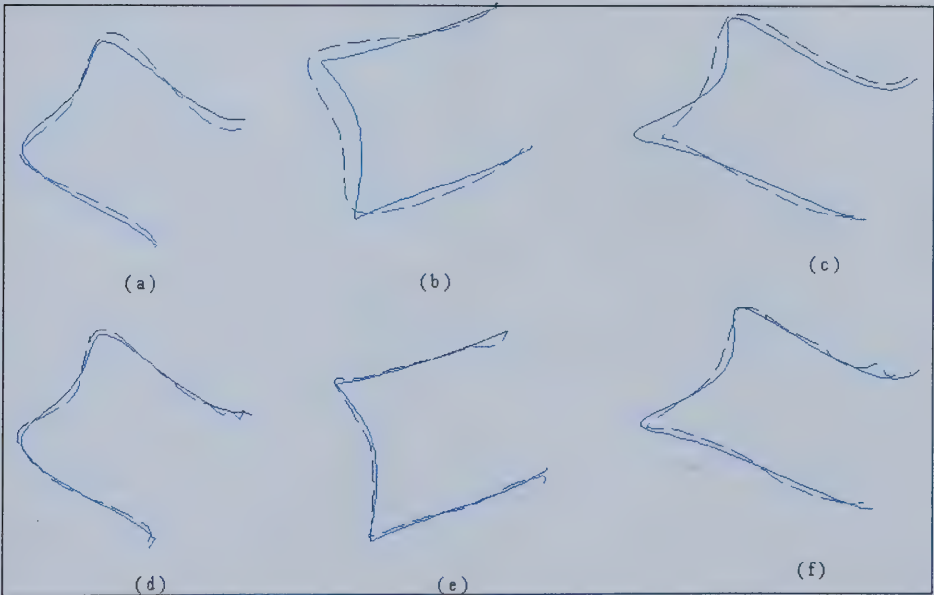


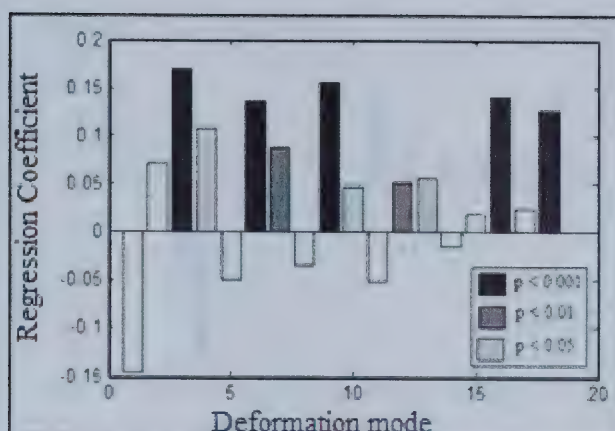
Figure 5.7 Shape prediction for normal and abnormal vertebrae by using PCA-based model [36] (top row), and the proposed model (bottom row). (a,d) are normal vertebrae, (b,c,e,f) are vertebrae with osteophytes pathology.

5.3.2 Evaluation of Wavelet-ICA shape analysis framework

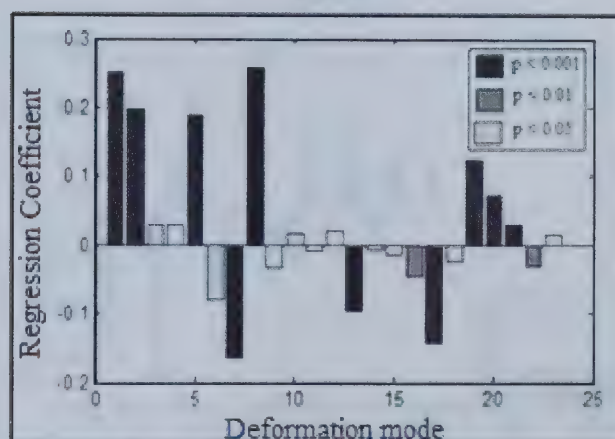
We evaluate the ability of the Wavelet-ICA shape analysis framework to explain group differences between normal and pathological vertebrae. We also compare the proposed analysis framework against a PCA-based shape analysis scheme.

In section 5.2, we use univariate regression analysis to relate the modes of deformation of the multi-scale shape model to a clinical dichromatic outcome variable. In our experiments, the clinical variable is the anterior osteophyte pathology. Figures 5.8-9 show the resulting regression coefficients of the localized deformation modes and the corresponding significance levels for the lumbar vertebrae L1-L5. For the vertebrae L1-L5, the percentages of localized deformation modes that have a statistically significant relationship with the clinical variable ($p < 0.05$) are as follows: 55 %, 58 %, 38.8 %, 65.52%, and 57.7 %, respectively.

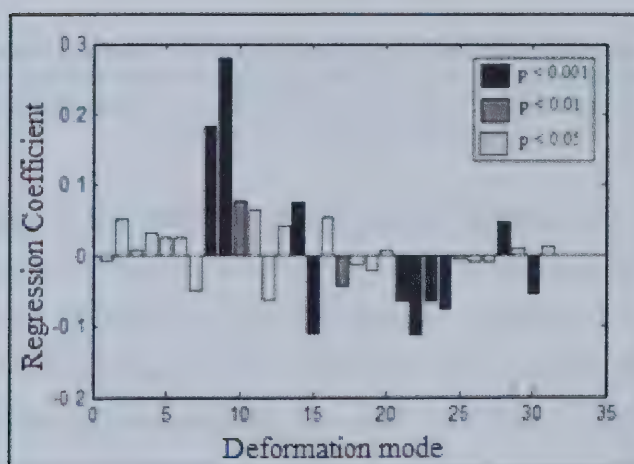
For all vertebrae investigated, an average of 55 % of the spatially localized deformation modes was found to have a statistically significant relationship with the osteophyte clinical variable. This finding suggests that the proposed localized, pathology-related shape parameters can be further exploited in applications such as pathology-based image retrieval.



(a) L1

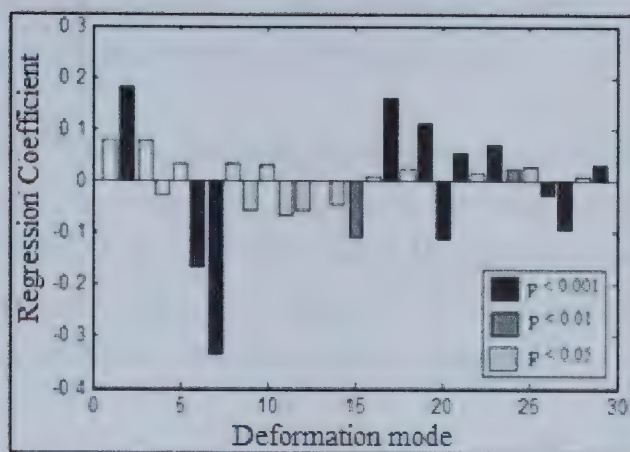


(b) L2

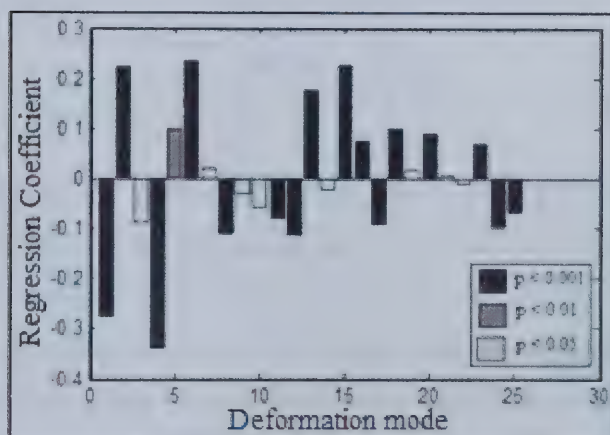


(c) L3

Figure 5.8 Regression coefficients from the investigation of the relationship between the localized deformation modes and the osteophyte pathology for lumbar vertebrae L1-L3. Statistical significance levels of the localized deformation modes are indicated by the color code: black ($p < 0.001$), dark gray ($p < 0.01$), and light gray ($p < 0.05$)



(a) L4



(b) L5

Figure 5.9 Regression coefficients from the investigation of the relationship between the localized deformation modes and the osteophyte pathology for lumbar vertebrae L4-L5. Statistical significance levels of the localized deformation modes are indicated by the color code: black ($p < 0.001$), dark gray ($p < 0.01$), and light gray ($p < 0.05$)

5.3.3 Evaluation of the multi-vertebrae classification scheme

The performance of the proposed classification scheme is evaluated by using percentages of the correct class classification, and the Receiver Operating Characteristic curves (ROC). The feature discriminatory power (FDP) criterion is used for the feature selection step. We compare the performance of the proposed classification scheme against a PCA-based classification scheme. We use a training set to construct a PCA-based basis of eigenshapes. These are similar to the basis functions of the Point Distribution Model (PDM) proposed by Cootes et al. [35]. The features are extracted by projecting the vertebral shapes onto the space of eigenshapes. The features are then ranked in

descending order of the eigenvalues. This classification scheme is similar to the shape-based classification method proposed by Roberts et al. [92].

We measure the ability of the proposed classification scheme to discriminate between normal and pathological vertebrae by using the True Positive Rate (TPR) and the True Negative Rate (TNR). A True Positive Rate is defined as the percentage of normal vertebrae that are correctly classified as normal. The True Negative Rate is the percentage of abnormal vertebrae that are correctly classified as abnormal:

$$TPR = \frac{TP}{TP + FN} \quad (5.22)$$

$$TNR = \frac{TN}{TN + FP} . \quad (5.23)$$

Table 5.1 summarizes the correct normal and abnormal classification rates for the vertebrae of the lumbar spine. The results are reported for the proposed classification scheme and the PCA-based scheme. The Fisher linear discriminant classifier is used in both frameworks. In comparison to the PCA-based classification scheme, the proposed method has achieved significant improvements in the classification rates. Table 5.1 shows improvements of 14.8 % and 53.60 % in the correct classification rates of normal and pathological L1 vertebra. The proposed scheme has achieved improvements of 35.37 %, and 33.50% in the correct classification rates of normal and pathological L2 vertebra. For the L3 vertebrae, the improvement percentages are 40.12% and 6.55% for normal and pathological classes, respectively. Improvement percentages of 29.74 % and 11.75 % are obtained in the correct classification rates of the L4 vertebra. For the L5 vertebrae, the improvement percentages are 18.28% and 16.82% for the normal and pathological cases, respectively.

For the entire lumbar spine L1-L5, the average and standard deviation classification rates of the proposed method are 96.08 % and 3.47 % for normal vertebrae, and 94.58 % and 4.82 % for abnormal vertebrae, respectively. In comparison to the PCA-based classification scheme, our method has a more consistent performance over all the vertebrae of the lumbar spine. The localized Wavelet-ICA features tend to better quantify the concavities in the vertebral body across the lumbar spine.

Table 5.1 Correct classification rates of the proposed classification scheme, and PCA-based classification scheme for detection of anterior osteophyte in lumbar vertebrae.

	Proposed classification method		PCA-based classification method	
Vertebra Index	correct normal % (TPR)	correct abnormal % (TNR)	correct normal % (TPR)	correct abnormal % (TNR)
L1	97.92	87.04	85.26	56.67
L2	98.28	98.83	72.60	74.03
L3	92.35	97.50	65.91	91.51
L4	92.31	96.92	71.15	86.73
L5	99.53	92.59	84.15	79.26
Mean	96.08	94.58	75.81	77.64
Standard deviation	3.47	4.82	8.50	13.51

Receiver operating characteristics (ROC) curves are also used to assess the ability of the proposed framework to discriminate between normal and pathological vertebrae. The ROC is a plot of the True Positive Rate (TPR) versus the False Positive Rate (FPR) as the threshold varies. In the statistical learning literature, The True Positive Rate (TPR) is also referred to as the sensitivity, and the False Positive Rate is defined as (1-specificity). The ROC curve provides a good method for visualizing the effect of changing the threshold on the classification accuracy. The Area under the ROC curve (AUC) is also used in the evaluation process. The AUC measures the seperability of the distributions of the normal and pathological classes.

Figures 5.10-14 show the ROC curves for the proposed classification scheme (WICA-based) and the PCA-based scheme [92] for the vertebrae of the lumbar spine L1-L5. The areas under the curves (AUC) for the ROC curves are given in Table 5.2. Table 5.3 compares the specefity of the two classification schemes at sensitivities of 0.80 and 0.90.

The proposed WICA-based classification scheme appears to dominate the ROC curves for the entire lumbar spine L1-L5. At a sensitivity of 80%, the proposed WICA-based classification scheme achieves an average specefity of 99.25 %, versus an average specefity of 72.9 % for the PCA-based classification scheme. At a sensitivity level of

90 %, the WICA-based classification scheme achieves an average specefity of 97.17 % versus 52.88 % for the PCA-based model.

At a False Positive Rate (FPR) of 0.2 (i.e., the specefity equals 80 %), the proposed classification scheme achieves an average sensitivity of 95.2 %, versus 78.2% for the PCA-based scheme. At a False Positive Rate (FPR) of 0.05 (i.e., the specefity equals 95 %), the proposed classification scheme achieves an average sensitivity of 93.2 %, versus 59.2% for the PCA-based scheme. In terms of the Area under the Curve (AUC), the proposed classification scheme has achieved an average improvement of 14.2 %.

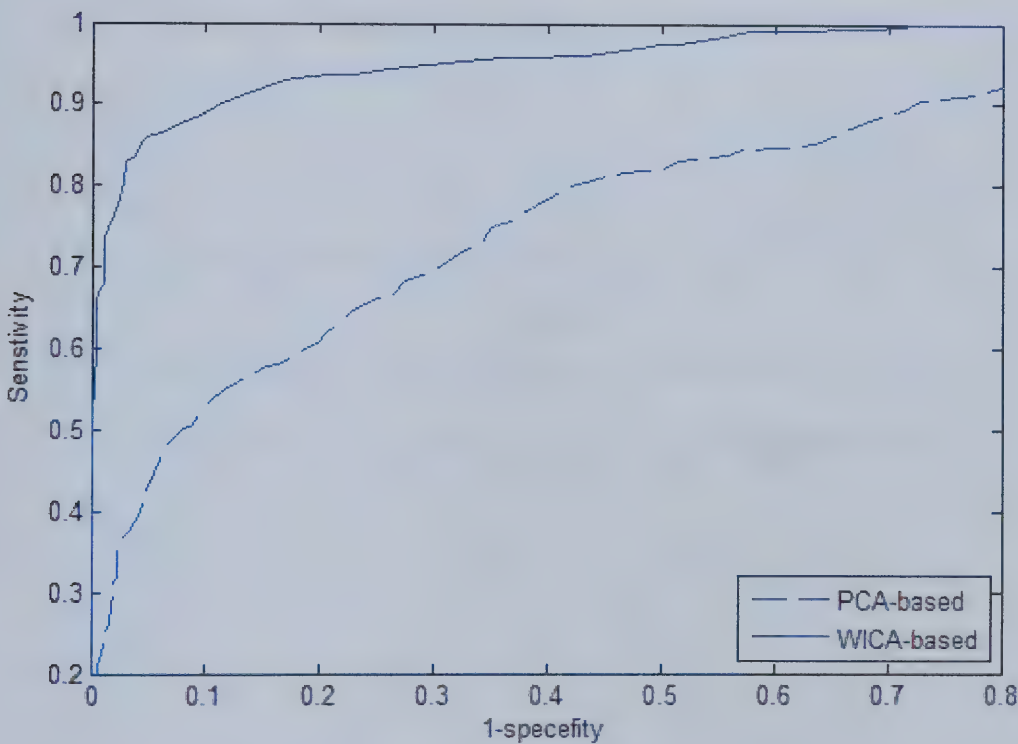


Figure 5.10 ROC curves for detection of anterior osteophytes in L1 vertebra

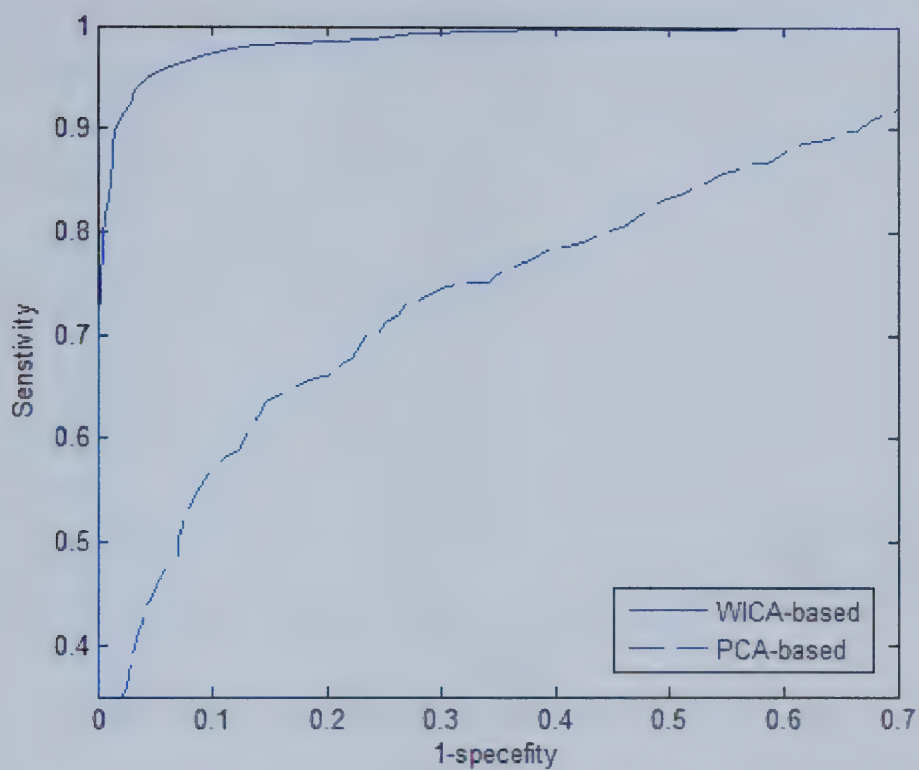


Figure 5.11 ROC curves for detection of anterior osteophytes in L2 vertebra

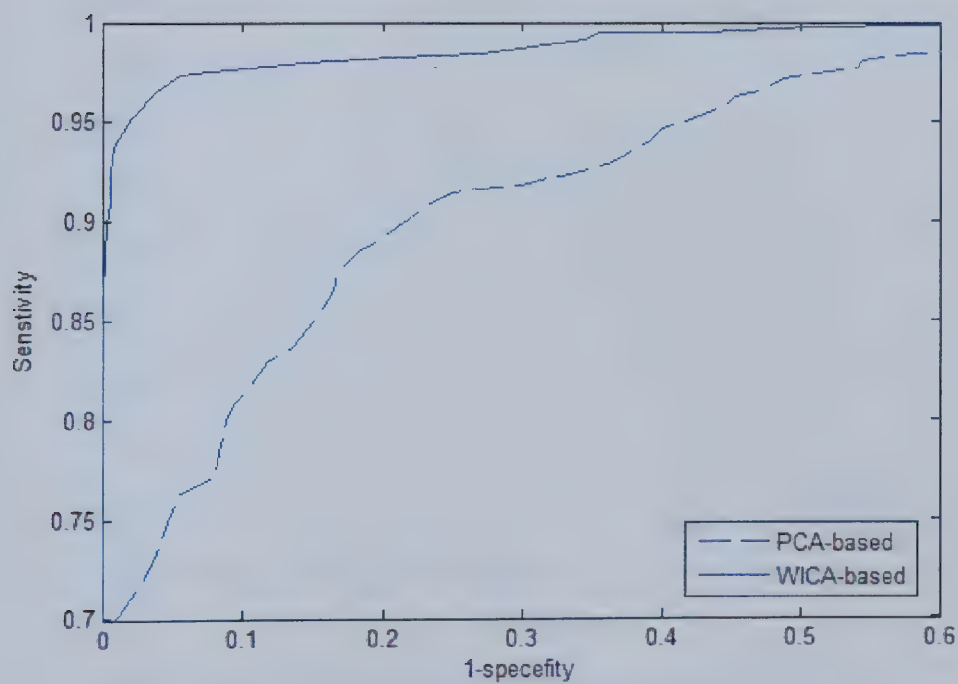


Figure 5.12 ROC curves for detection of anterior osteophytes in L3 vertebra

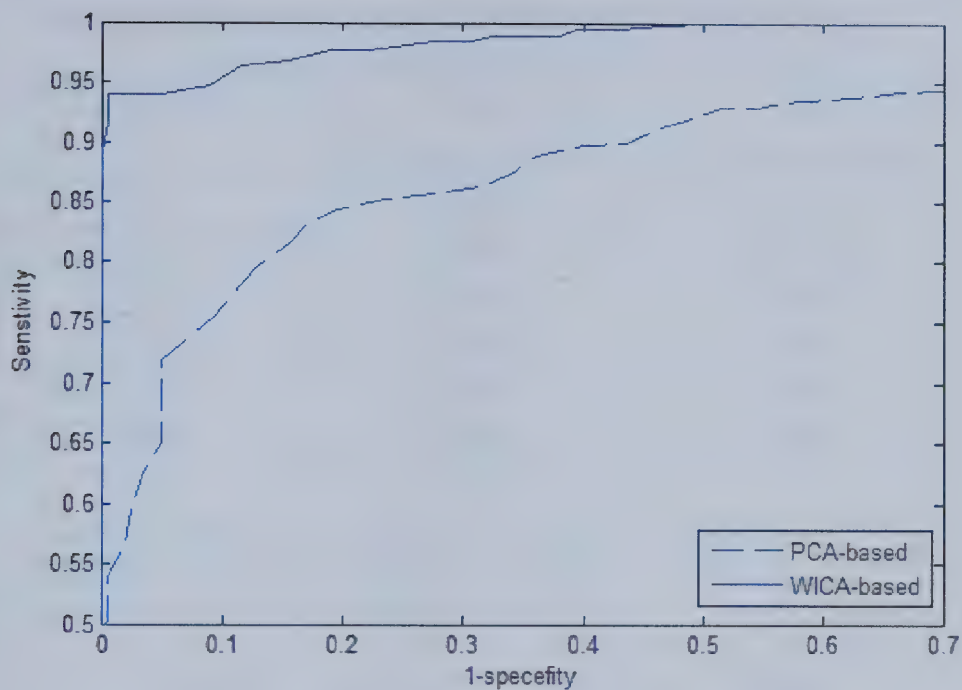


Figure 5.13 ROC curves for detection of anterior osteophytes in L4 vertebra

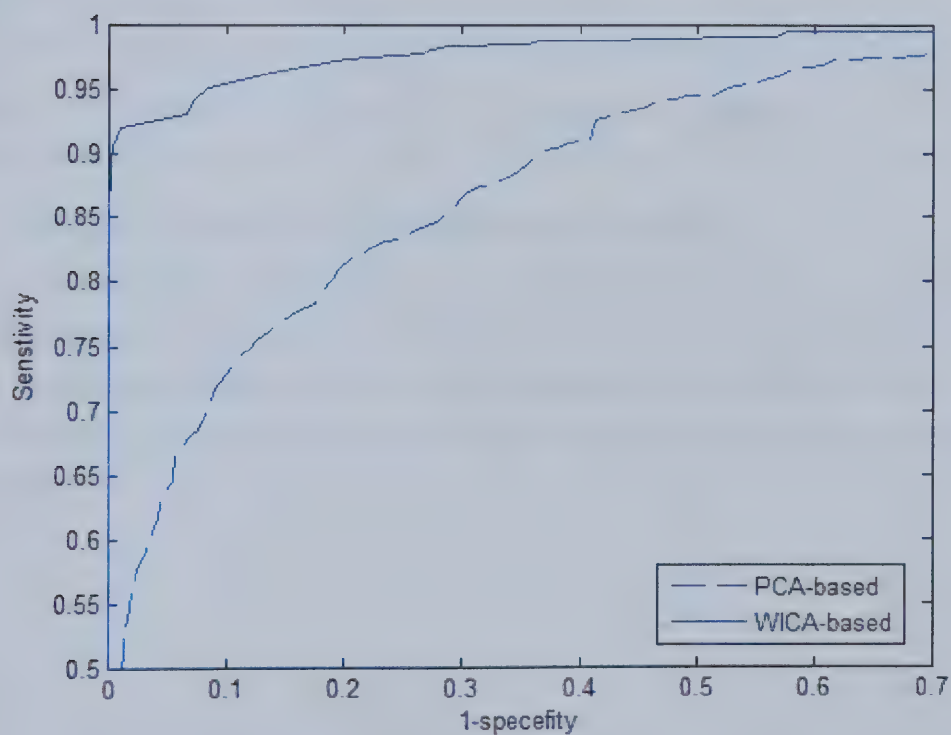


Figure 5.14 ROC curves for detection of anterior osteophytes in L5 vertebra

Table 5.2 Areas under curves for ROC curves of the proposed classification scheme and the PCA-based classification scheme.

Vertebra Index	AUC	AUC
	WICA-based classification	PCA-based classification
L1	0.9535	0.7701
L2	0.9876	0.8015
L3	0.9811	0.9357
L4	0.9833	0.8797
L5	0.9826	0.8933
Mean	0.9776	0.8561

Table 5.3 Specefity at various sensitivities for the Lumbar vertebrae L1-L5 .

Vertebra	Specefity at 80 % sensitivity		Specefity at 90 % sensitivity	
	WICA-based	PCA-based	WICA-based	PCA-based
L1	97.45 %	57.5 %	88.5 %	30%
L2	98.8 %	56 %	97.8 %	34 %
L3	100 %	92 %	99.74 %	77%
L4	100 %	86.4 %	99.83 %	60%
L5	100%	72.6 %	99.98 %	63.4%

5.3.2 Effect of feature selection on classification accuracy

This section investigates the effect of the feature selection stage on the overall classification accuracy. We evaluate the performance of two feature selection criteria: the feature spatial locality (FSL) and the feature discriminatory power (FDP). Table 5.4 shows the average correct classification rates achieved by using different feature selection approaches.

Table 5.4 Effect of various feature selection approaches on the class classification accuracy.

Feature selection Method	correct classification rate of normal vertebra	correct classification rate of abnormal vertebra
No feature selection	87.83 %	83.46 %
Using spatial locality criterion only	94.966 %	91.776 %
Using discriminatory power criterion only	96.08 %	94.58 %
Fusion of the two selection criteria	97.528 %	94.522 %

Based upon the figures in table 5.4, the following comments can be made:

- In general, the feature selection step has significantly improved the classification rates of normal and abnormal classes.
- By using the feature spatial locality criterion (FSL), we achieve an increase of 7.14 %, and 8.32 % in the correct classification rates of normal and abnormal vertebra, respectively.
- By using the feature discriminatory power criterion (FDP), we achieve an increase of 8.25% and 11.12% in the correct classification rates of normal and abnormal vertebra, respectively.
- A limited performance gain is achieved by the integration of the two selection criteria.

5.4 Discussion

The experiments in section 5.3 have demonstrated the ability of the proposed Wavelet-ICA shape analysis and classification framework to capture subtle discriminating features of the vertebral bodies. The proposed classification scheme has achieved an average classification rate of 97.53 % for normal vertebrae, and 94.52 % for vertebrae with anterior osteophyte pathology. These results compare well to those from related vertebral classification methods in the literature. In [120,121], researchers from the National

Library of Medicine presented classification schemes for lumbar and cervical vertebrae against anterior osteophytes. In [120], the authors reported average correct classification rates of 90.46 for normal lumbar vertebrae, and 89.48% for abnormal vertebrae or those with anterior osteophyte pathology. In [121], the authors reported classification accuracies of 84.44 % for normal cervical vertebrae, and 77.08 % for pathological cervical vertebrae. Moreover, we are able to perform multi-vertebrae classification due to the spatial locality of the Wavelet-ICA features. The sparsity and spatial locality of the Wavelet-ICA representation have enabled us to extract subsets of the feature space that are linked to each of the five lumbar vertebrae (L1-L5). This result allowed us to perform accurate multi-vertebrae classification for the complete lumbar spine.

We have also investigated the potential gains of the feature selection step on the Wavelet-ICA domain prior to classification. In general, the feature selection step has significantly improved the overall classification accuracy for both normal and abnormal vertebrae. We explored two criteria for selecting feature subsets: the feature spatial locality (FSL) and the feature discriminatory power (FDP). The sequential fusion of the two criteria has resulted in an increase of 9.39 % and 11.52% in the correct classification rates of normal and abnormal vertebrae, respectively. Better results were achieved by using the feature discriminatory power (FDP) criterion than by using the feature spatial locality criterion (FSL). This difference is attributed to the fact that the feature discriminatory power criterion is a more general criterion that allows us to select significant spatially localized features, as well as other discriminating features from neighbouring vertebrae.

5.5 Conclusion

In this chapter, we investigated the potential of the novel multi-scale shape model in the quantification and diagnosis of vertebral deformities.

We presented a novel vertebral shape analysis and classification framework using localized, pathology-related Wavelet-ICA parameters. Within a Wavelet-ICA regression analysis framework, we were able to establish a statistically-relevant relationship between the shape parameters and clinical information. Our experiments demonstrated the ability of the proposed framework to discriminate between classes of normal and pathological vertebrae. This ability opens the door for other anatomically meaningful interpretations of the modes of deformations of the multi-scale shape model. Furthermore, we presented a

novel method for computer-assisted multi-vertebrae classification, which incorporates the following main steps: localized shape modeling, feature selection, and statistical classification. High classification accuracies were achieved in the experiments. Thanks to the locality of the features, the new classification scheme can simultaneously assess the condition of the complete lumbar spine against the pathology of interest.

6. Prior-driven Segmentation Using Local Appearance Profiles in Contourlet Domain and Multi-scale Shape Model

6.1. Introduction

In this chapter, we attempt to bring recent advances in the fields of image processing and computational harmonic analysis to the field of medical imaging. Within this context, we propose novel methods for the segmentation and retrieval of anatomical structures by using Contourlet-based local appearance profiles and the multi-scale shape prior.

The segmentation of anatomical structures in medical images is an essential step for various medical applications such as computer-aided diagnosis, 3D visualization, and surgical planning. Manual segmentation methods are both time-consuming and non-reproducible. Hence, developing efficient computer-assisted segmentation methods has been a long-standing goal within the medical imaging research community. In this work, we are particularly interested in the segmentation of human vertebrae in digital radiographs. Accurate vertebral segmentation plays a vital role in the proper assessment of various vertebral abnormalities. Nevertheless, efficient vertebrae segmentation continues to be a challenging task for a number of reasons. First, poor image quality and noisy and incomplete data tend to affect the ability of low-level methods to accurately detect the boundary of the vertebral body. Second, the complexity of the anatomical structure of the human spine and the wide range of possible deformations tend to limit the ability of top-level prior-based methods to generalize well to unseen examples. Thus, an efficient vertebrae segmentation algorithm needs to keep the balance between the two extremes. On the one hand, it should have enough prior knowledge to reduce the sensitivity to noise and image quality. On the other hand, it should still be able to detect valid deformations.

Another medical application of interest in this chapter is content-based image retrieval (CBIR), which refers to the retrieval of images in the database by using visual features extracted from the images, rather than textual information. In the past few years, CBIR has received increasing attention within the research community for two main reasons:

(i) the recent advances in digital imaging technologies and data compression methods, and (ii) the substantial growth in the number of digital images in hospitals and the

increasing popularity of Picture Archival & Communications systems (PACs). Image retrieval systems have the potential to serve many educational and diagnostic purposes within the medical imaging community.

For two major research groups, the segmentation and retrieval of spine radiographs have been the main focus. The first group is the ListerHill National center of Biomedical Communications, which is an R&D division of the US National library of Medicine (NLM). The NLM has made available a set of 1,700 cervical and lumbar spine digital radiographs with various pathologies that were collected in the second National Health and Nutrition Examination Survey (NANESII) conducted by the US National Center for Health Statistics (NCHS) [108]. These images have been used by various researchers worldwide who are interested in the segmentation and retrieval of spine x-rays. Another major effort within the field is a collaborative project carried out by the departments of Radiology, Medical Information, and Computing Science at the Aachen University of Technology in Germany. The Image Retrieval for Medical Applications (IRMA) project aims at developing content-based image retrieval systems for diagnostic tasks. A large database of radiographs has been made available to the research community [125].

The rest of this chapter is organized as follows: section 6.2 highlights the main contributions presented in this chapter. Section 6.3 presents the details for a novel appearance-based salient point detection and matching algorithm. Section 6.4 presents a novel vertebrae segmentation method. Section 6.5 presents the details for a novel Content-based medical image retrieval system. Our experimental results are presented in section 6.6, and our conclusions are drawn in section 6.7.

6.2. Contributions

Figure 6.1 illustrates the general framework for the methods presented in this chapter.

The following outlines the main contributions presented in this chapter:

- In an attempt to bring recent advances in image processing and computational harmonic analysis to the field of medical imaging, we present a novel salient point detection and matching algorithm based upon Non-Subsampled Contourlet transform (NSCT), a truly 2D multi-scale directional image analysis tool. In contrast to low-level salient point detection methods, our method integrates prior knowledge about the local appearance of the salient points. The proposed method has proven successful in detecting a reduced number of robust and stable salient points that are most relevant to the structure of interest. The Contourlet-based salient point detection method is central in the new methods for the segmentation and retrieval of radiographs.
- In an attempt to overcome challenges in the segmentation of vertebral structures, we present a novel segmentation method using shape-driven salient point matching in the contourlet domain. At the low-level, the method uses the Non Subsampled Contourlet Transform (NSCT) to detect robust and repeatable salient points in the spine x-ray. Rather than using global texture models to drive the segmentation process, we construct contourlet-based local appearance profiles at salient points of the vertebral bodies. A multi-scale shape prior is also constructed as described in chapter four. During the testing phase, the contourlet-based local appearance model is used to detect the vertebral boundaries in the test x-ray image, while the localized multi-scale shape prior is used to drive the segmentation process. The new method is shown to have high performance in terms of robustness to noisy images and generalization to unseen cases.
- We have also used the proposed contourlet-based salient point detection method in a content-based image retrieval task. The new method uses the robustness and stability of the extracted salient points, and the information-rich Contourlet descriptors to retrieve medical images of similar classes.

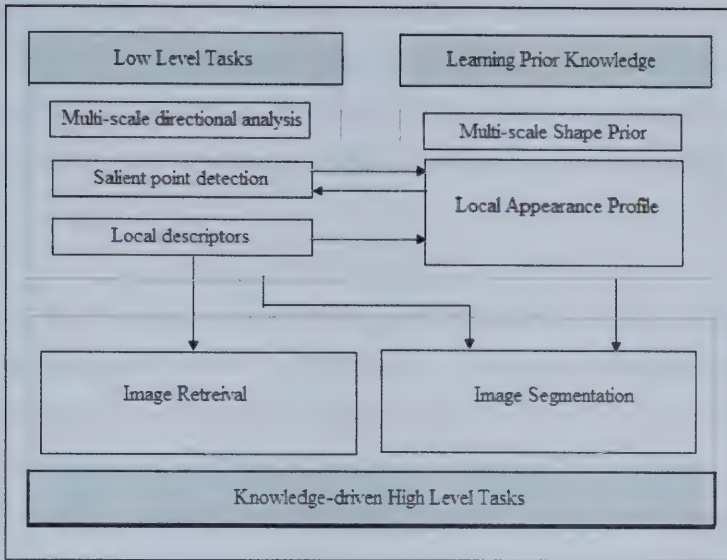


Figure 6.1 General framework for knowledge-driven methods presented in this chapter

6.3 Multi-scale salient point detection in Contourlet domain

This section presents the details for a new method for detecting salient points in medical images based upon Non Subsampled Contourlet analysis. This method is central to the segmentation and retrieval methods presented in sections 6.4 and 6.5, respectively.

6.3.1 Background

Based upon our understanding of the Human Visual System (HVS), it is believed that the human vision tends to be selectively attentive to localized “different” regions of the visual field. The detection of these salient regions enables the brain to perform other higher-level visual tasks such as pattern recognition and object tracking. In the literature, these regions of interests are often referred to as interest points, key points, or salient points. No exact definition exists for what makes a specific region be considered as a salient region. However, the expression is generally used to refer to regions of the visual field with localized distinctive features such as edges, corners and blobs. All these are thought to provide cues for recognizing shapes and objects in complex visual scenes.

Inspired by the Human Visual System (HVS), salient point detection methods have been used as a low-level image processing procedure in many computer vision tasks including registration, tracking, robot navigation, and medical imaging. The performance of the salient point detection algorithm has a great impact on subsequent higher-level tasks.

Generally speaking, salient points need to be well defined and stable under different transformations. Various salient point detection algorithms are often evaluated in terms of repeatability, which is defined as the ability of the method to detect the same salient points under various transformations. Popular transformations are scaling, rotation, and change in view.

Several salient point detection approaches were proposed in the literature. In [126], Schmidt et al. grouped salient point detection methods according to the theory of operation into three categories: edge-based, intensity-based, and biologically-inspired methods. One of the most famous salient point detectors is the Harris corner detector [127], which detects corners in an image by using the autocorrelation matrix. Despite being useful in relatively simple scenes, the Harris detector suffers from a number of limitations. First, it is based on single scale analysis. Second, it suffers from poor localization ability. The SIFT algorithm is another important salient point detection method [128]. It extracts salient points by finding the local maxima of the response of the Difference-of-Gaussian (DOG) operator. Biologically inspired salient point detectors focus on directional and multi-scale analysis of images. Tian [129] and Loupais et al. [130] proposed wavelet-based salient point detectors. The wavelet transform has a number of appealing properties such as multi-scale representation and spatial-frequency localization. Nevertheless, wavelets suffer from a lack of shift invariance and limited orientation selectivity, both are important properties for the efficient detection of salient points in complex scenes such as medical images. In [131], Itti et al. proposed a biologically inspired bottom-up saliency detector based upon Gabor filters. In [132], the authors presented a salient point detector based upon the phase congruency of Log-Gabor filter response. In [133], Gao et al. proposed a multi-scale corner detection method based upon the Log-Gabor transform. In comparison to wavelets, Gabor filters tend to offer better directional and multi-scale analysis.

6.3.2 Methodology

In this chapter, we are interested in salient point detection and matching in medical images. Detecting robust and stable salient points of interest in medical images is a challenging task, because of the low quality of the images, and the complexity of the anatomical structures of interest.

In medical image analysis, an ideal salient point detector needs to exhibit the following characteristics:

- The detector should be able to extract information from multiple scales and orientations. No strict definition exists for what constitutes a salient point in an image. Thus, it is logical to use biologically inspired filters in the detection process.
- The localization property of the detector is also an important aspect that tends to affect other higher level tasks such as feature extraction, matching, and registration.
- Also, the detector should be able to capture higher-level salient structures in the image (e.g., contours).
- The computational aspect of the detector is important since salient point detection always occurs at the early stages of the computer vision system.
- In most medical image analysis applications, we are interested in a particular object. Thus, it is beneficial for the salient point detector to be target-driven. It can be made so by integrating prior knowledge to the salient point detection procedure.

In an attempt to achieve some of these features, we present a new method for salient point detection that is suitable for the extraction of robust points of interest in medical images. The new method is utilized as an early image processing stage that serves the other higher-level imaging methods presented in section 6.4. In this section, we propose a new appearance-based salient point detection method based upon Non Sub-sampled Contourlet Transform (NSCT). The multi-scale and directional capabilities of the NSCT transform are utilized to extract salient points that are robust to noise and rigid and non-rigid deformations. Moreover, the proposed method also utilizes prior knowledge about the local appearance of the object of interest. This ability allows for the detection of a reduced set of salient points that are most relevant to the structure of interest.

6.3.2.1 Salient point detector

This section presents the details for a novel salient point detection method using Non-Subsampled Contourlet Transform (NSCT). Using the capabilities of the NSCT transform to capture higher-order image structures, we are able to extract robust salient points in medical images. In the context of our work, salient points refer to significant corners and

smooth contours that can be used as cues during the segmentation and retrieval processes. The new method is based upon the NSCT and the second moment matrix. We will refer to the new method as NSCT-SMM. The input image is first decomposed by using the NSCT transform at different scales and orientations. The magnitudes of the decomposition at all scales and orientations are then formulated into a second moment matrix. Following classical moment analysis, [134,135], intensity variation in the contourlet-based feature image is characterized by the two eigen axes of the second moment matrix. The strengths of the maximum and minimum moments are used to infer the existence of significant corners and contours, respectively. Figure 6.2 shows a block diagram for the proposed method.

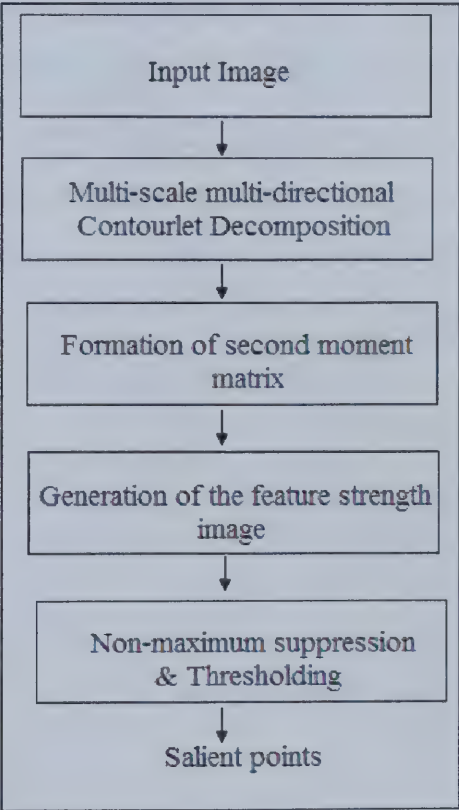


Figure 6.2 Flowchart of the proposed multi-scale contourlet-based salient point detector

The following explains the steps for the proposed method:

a. Non Sub-sampled Contourlet decomposition: A multi-scale directional decomposition is achieved by using the NSCT transform. The input image $I(x, y)$ is decomposed into 4 scales. In the first scale level, we use a 4-level Directional Filter Bank (DFB) to get 16 directional frequency partitioning. In the second and third scales, eight directional frequency partitions are obtained by using a 3-level Directional Filter bank. In the final scale level, 2-level directional filter-banks are used. Let $\{B_j\}, j = 0, 1, 2, 3$ denote the set of response images that results from the non sub-sampled pyramid decomposition step, where j is the scale level. At each scale j , the response image B_j is further decomposed by using a l_j -level non sub-sampled Directional Filter Bank to produce a set of multi-scale multi-directional contourlet sub-band images $\{W_{j,k}\}$, where $k = 0, 1, \dots, 2^{l_j-1}$, and $j = 0, 1, 2, 3$. Figure 6.3 shows the oriented contourlet basis functions for two different scales. Due to the flexibility of the transform, we can decompose the image by using different angular resolutions at every scale.

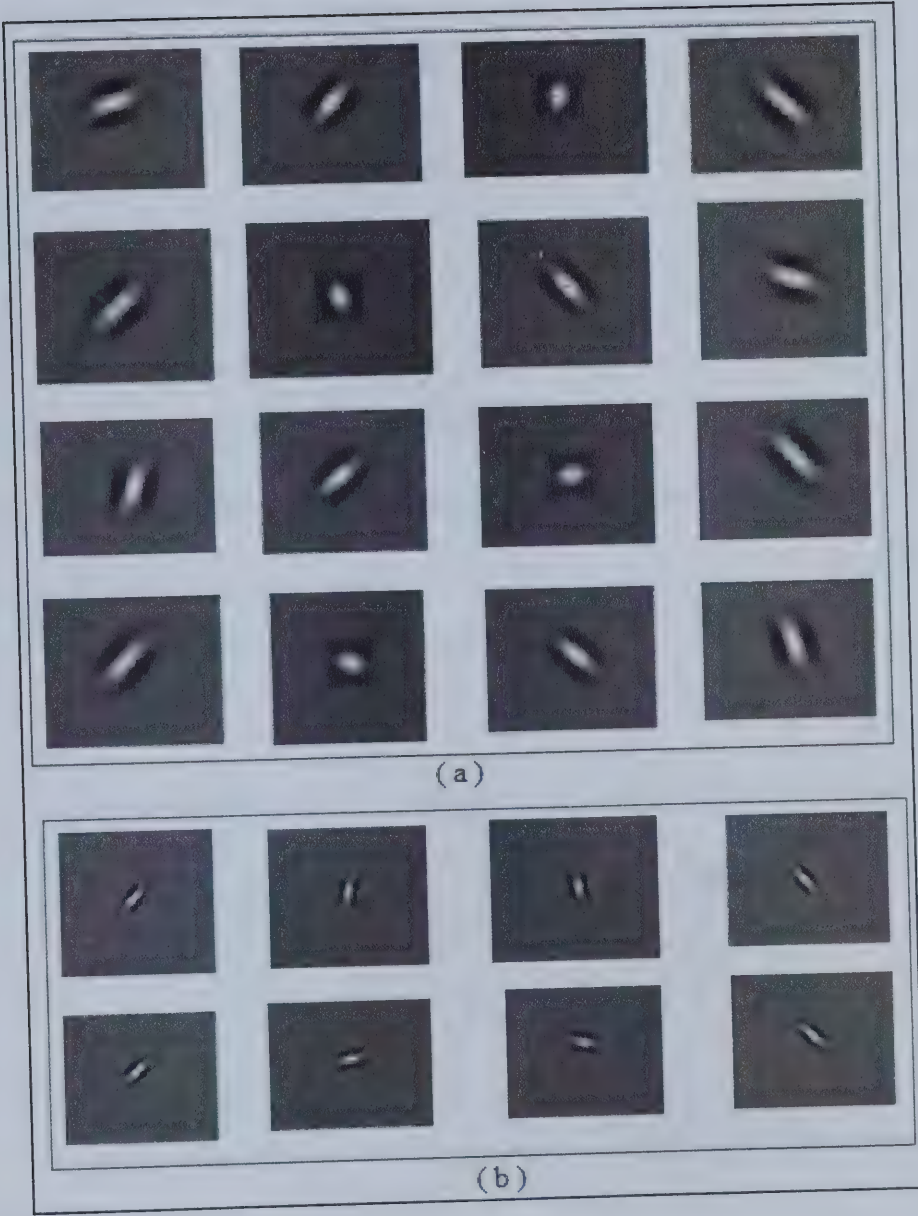


Figure 6.3 NSCT basis functions for two scales with different angular resolutions at each scale

b. Following classical moment analysis equations, [134,135], we compute the following:

$$a = \sum_{j=0}^M \sum_{k=0}^{N_j} (|W_{j,k}| \cos(\theta_k))^2 \quad (6.1)$$

$$b = \sum_{j=0}^M \sum_{k=0}^{N_j} |W_{j,k}| \cos(\theta_k) \sin(\theta_k) \quad (6.2)$$

$$c = \sum_{j=0}^M \sum_{k=0}^{N_j} (|W_{j,k}| \sin(\theta_k))^2, \quad (6.3)$$

where

j is the scale of the decomposition.

M is the maximum number of scales.

k is the orientation at scale j .

N_j is the number of decomposition orientations at scale j .

$|W_{j,k}|$ is the magnitude of the non subsampled contourlet coefficients at scale j and orientation k .

$\theta_{j,k} = \frac{(k-1)\pi}{N_j}$ is the angle of the k^{th} orientation at the j^{th} scale level.

c. The second moment matrix is then constructed in the contourlet domain as

$$M = \begin{bmatrix} a & b \\ b & c \end{bmatrix}. \quad (6.4)$$

d. The eigenvalues of the second moment matrix in equation 6.4 is then calculated as follows:

$$E_1 = \frac{1}{2}[a+c] - \frac{1}{2}\sqrt{4b^2 + (a-c)^2} \quad (6.5)$$

$$E_2 = \frac{1}{2}[a+c] + \frac{1}{2}\sqrt{4b^2 + (a-c)^2}. \quad (6.6)$$

The eigenvalue decomposition of the second moment matrix M provides a description of the gradient structure in the image. The calculated eigenvalues E_1 and E_2 indicate the certainty of along the directions of their associated eigenvectors [134].

e. Compute the Contourlet-based feature strength image: Our main goal is to use the salient point detector for the segmentation of the anatomical structures in x-ray images. We would like to extract salient points which correspond to the significant corners and smooth contours of the object of interest. Thus, we define a contourlet-based feature strength image (F) as a weighted sum of the two eigenvalues of the second moment matrix. The smaller eigenvalue (E_1) is a measure of the corners in the image, and the larger eigenvalue (E_2) is a measure of the contours and edges in the image [134]:

$$F = \alpha E_1 + (1 - \alpha) E_2 . \tag{6.7}$$

Figure 6.4 shows an example of x-ray image of the lumbar spine and the corresponding Contourlet –based feature strength map.

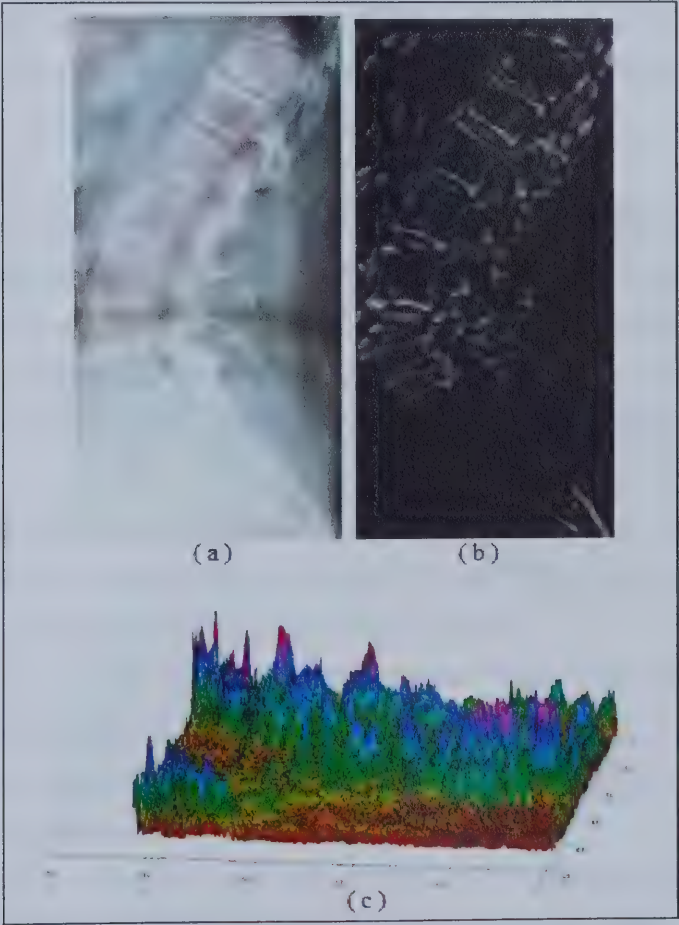


Figure 6.4 Example x-ray image and corresponding Contourlet-based feature strength image.

f. Non-maximum suppression and thresholding: Salient points are finally extracted from the Contourlet-based feature image by using non-maximum suppression and thresholding. This step selects local maxima points as the salient points.

6.3.2.2 Incorporation of local appearance knowledge

As discussed before, salient point detectors usually appear at early stages of computer vision tasks. Thus, the number of detected points greatly affects the efficiency of the whole application. Most of the methods in the literature rely on low-level operations where no prior knowledge is incorporated in the detection process. However, in many cases, we have prior knowledge about the object of interest. Thus, our goal is to perform a sort of object-guided search for the salient point, rather than a pure bottom-up search.

We propose to further enhance the proposed salient point detection method described in section 6.3.2.1 by incorporating prior knowledge about the object of interest. Doing so has the advantage of reducing the number of detected salient points for subsequent complex tasks. This procedure will increase the efficiency of the overall computer vision task at hand (i.e. segmentation).

The new method benefits from the rich multi-scale directional information offered by the Non Subsampled Contourlet decomposition to create localized descriptors for the salient points of the object of interest. The local descriptors are then modeled by using Mixture of Gaussians to construct a local appearance model of the target salient points. The following summarizes the steps for the new appearance-based salient point detection method:

1. We start with a training set of x-ray images with a number of ground truth salient points for the object of interest. For each ground truth salient point, we extract a local appearance profile based upon NSCT features.
2. Using the set of NSCT descriptors, we construct local appearance profiles for the target salient points by using Mixture of Gaussians.
3. For a new image, we extract the candidate salient points by using the bottom-up approach explained in section 6.3.2.1.
4. A reduced set of salient points is then obtained by estimating the likelihood that every detected point corresponds to the learnt appearance model.

The following presents the details for the proposed method:

i. Feature Extraction

Given a set of M training images with a ground truth segmentation for the object of interest, we extract a set of N ground truth salient points from each image. First, we use the method described in section 6.3.2.1 to detect the salient points in the image. Next, ground truth salient points that correspond to the object of interest are extracted by using the ground truth segmentation. A ground truth salient point is defined as a salient point that has a point-to-curve distance of less than 5 pixels. Figure 6.5 shows two x-ray images with the ground truth salient points that define the border of the lumbar vertebrae L1-L5.

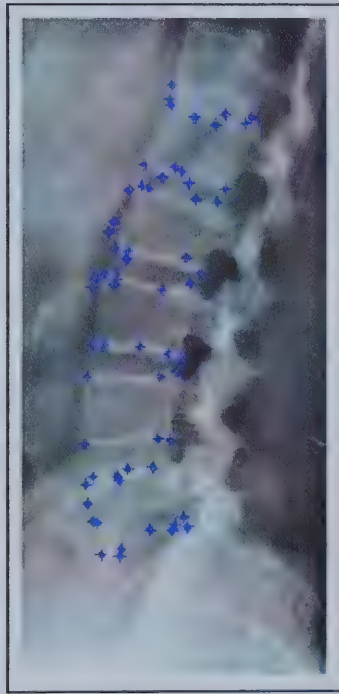


Figure 6.5 Example of ground truth salient points that define the lumbar spine in the x-ray images

For every ground truth salient point (P_i), we extract a local descriptor based upon the Non Subsampled Contourlet decomposition. The local descriptor is constructed by using the NSCT feature images obtained in section 6.3.2.1. We use a square grid of $N \times N$ ($N=7$) points. The ground truth salient point is at the center of the grid. Variance-based features are then extracted from the NSCT directional sub-bands at every scale, excluding the low-pass response.

For every NSCT feature image, we calculate the average absolute deviation of the gray values of the square grid centered at the salient point P_i :

$$E_j^i(a,b) = \frac{1}{N^2} \sum_{l=a-\frac{N}{2}}^{a+\frac{N}{2}} \sum_{k=b-\frac{N}{2}}^{b+\frac{N}{2}} |x_j^i(l,k) - \mu_j^i|, \quad (6.8)$$

where x_j denotes the j^{th} NSCT feature image, $x_j^i(l,k)$ is a coefficient at location (l,k) of the NSCT feature image, and μ_j^i is the mean of the coefficients in the $N \times N$ square grid whose center is the salient point P_i .

The contourlet-based local descriptor for the salient point P_i is then defined as

$$F_i = [E_1^i, E_2^i, \dots, E_j^i, \dots, E_D^i]^T, \quad (6.9)$$

where E_j^i is the average absolute deviation for the j^{th} NSCT sub-band, and D is the total number of NSCT sub-bands.

Figure 6.6 summarizes the complete procedure for the construction of the Contourlet-based local descriptor. The outcome of this stage is a set of $K=N \times M$ Contourlet-based feature vectors extracted from all the ground truth salient points in the training set. Next, we use Mixture of Gaussians to construct a local appearance profile for the salient points of the object of interest.

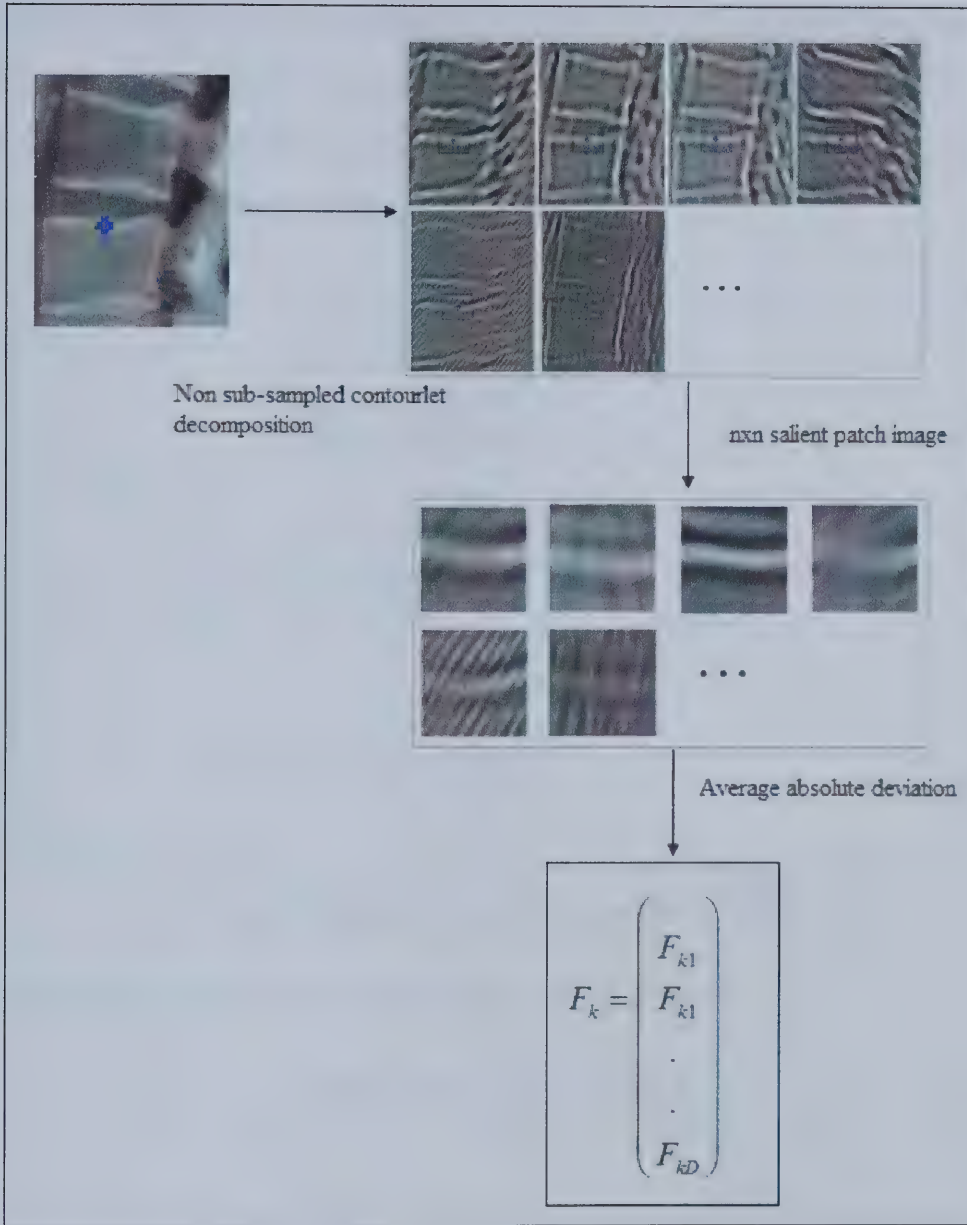


Figure 6.6 Extraction of Contourlet-based local descriptor for a salient point.

ii. Mixture of Gaussians Model

Let $\{F_k\}_{k=1..K}$ be the set of Contourlet-based local descriptors for the ground truth salient points in the training set. K is the total number of ground truth salient points, and F_k is the local descriptor of the k^{th} salient point.

We model the set of local descriptors as a mixture of N_m Gaussians. The Gaussian mixture model is a flexible parametric distribution. When N_m is large enough, any distribution can be modelled as a sum of N_m Gaussian functions [119]. Figure 6.7 shows the building blocks of the Gaussian Mixture model.

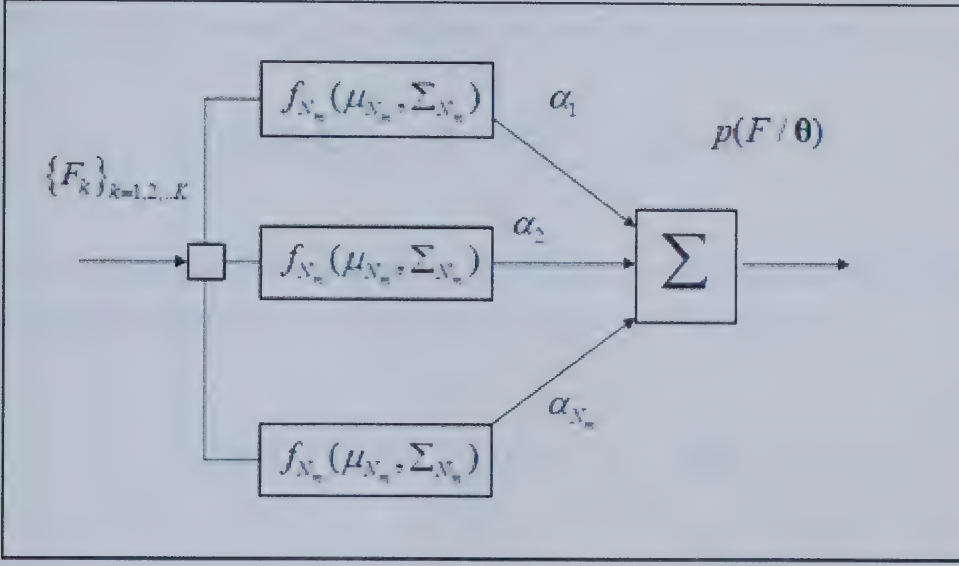


Figure 6.7 Building blocks of the Gaussian Mixture Model

The Gaussian Mixture model for the local descriptors is given by

$$p(F_k / \theta) = \sum_{n=1}^{N_m} \alpha_n f_n(F_k), \quad (6.10)$$

where N_m : Number of Gaussian functions in the model. In our model, we have $N_m=4$.

$f_n(F_k)$: Multivariate Gaussian density functions characterized by the mean (μ_n)

and covariance matrix (Σ_n):

$$f_n(F) = \frac{1}{(2\pi)^D (\Sigma_k)^{0.5}} \exp\left\{-\frac{1}{2}(F - \mu_n)^T (\Sigma_n)^{-1} (F - \mu_n)\right\}. \quad (6.11)$$

α_n is the mixture of weights, where $\sum_{n=1}^{N_m} \alpha_k = 1$.

The parameters of the Gaussian Mixture model are as follows:

$$\boldsymbol{\theta} = \{\alpha_n, \mu_n, \sum_n\}, \quad n = 1..N_m. \quad (6.12)$$

Using the training set $\{F_k\}_{k=1..K}$, we estimate the parameters of the Gaussian Mixture model (GMM) by maximizing the likelihood function. The GMM likelihood function is given by

$$p(\{F_k\}_{k=1..K} / \boldsymbol{\theta}) = \prod_{k=1}^K p(F_k / \boldsymbol{\theta}). \quad (6.13)$$

We use the Expectation-Maximization algorithm (EM-algorithm) to estimate the model parameters that maximize the likelihood function given in equation 6.13:

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \{p(\{F_k\}_{k=1..K} / \boldsymbol{\theta})\}. \quad (6.14)$$

The E-M algorithm estimates the parameters that maximize the likelihood function by iterating through two main steps: Expectation and Maximization. The model parameters are iteratively updated in the following manner:

E-Step:

Using the current estimate for the model parameters, we evaluate the posterior probabilities of the n^{th} Gaussian functions given the k^{th} local descriptor:

$$p(n / F_k) = \frac{\alpha_n f_n(F_k)}{\sum_{n=1}^{N_m} \alpha_n f_n(F_k)}. \quad (6.15)$$

M-Step:

The parameters of the Gaussian model are iteratively updated in the following manner:

$$\hat{\alpha}_n = \frac{1}{K} \sum_{k=1}^K p(n / F_k) \quad (6.16)$$

$$\hat{\mu}_n = \frac{\sum_{k=1}^K F_k p(n / F_k)}{\sum_{k=1}^K p(n / F_k)} \quad (6.17)$$

$$\hat{\sum}_n = \frac{\sum_{k=1}^K p(n / F_k) (F_k - \hat{\mu}_n)^2}{\sum_{k=1}^K p(n / F_k)}. \quad (6.18)$$

iii. Appearance-based salient point detection

The set $\boldsymbol{\theta} = \{\alpha_n, \mu_n, \sum_n\}$, $n = 1..N_m$ represents the estimated parameters for the local appearance model of the ground truth salient points. This model is used to bias the low-level salient point detection procedure described in section 6.3.2.1. For every detected low-level salient point, we extract a contourlet-based local descriptor by using the method shown in Figure 6.6. Next, we estimate the confidence level that the detected salient point actually belongs to the local appearance model learnt during the training phase. Following the Bayes decision rule [119], we use the posterior probability as the measure for the confidence level.

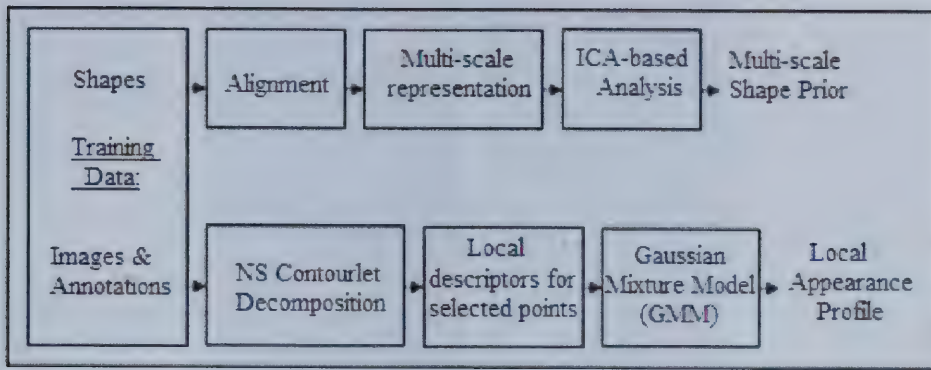
6.4 Vertebrae segmentation using contourlet-based salient point matching and multi-scale shape prior

This section presents the details for a new method for vertebrae segmentation in x-ray images. The human spine is a complex articulated structure exhibiting local and global deformations. In an attempt to overcome the challenges in the segmentation of vertebral structures, we present a novel segmentation method using shape-driven salient point matching in the Contourlet domain. The new method incorporates the contourlet-based salient point detection method and the multi-scale shape prior presented in chapter four. At the low level, the method uses the Non Sub-sampled Contourlet Transform (NSCT) as the multi-scale and multi-directional analysis tool. Using the method described in section 6.3, we are able to extract robust and repeatable salient points related to the lumbar spine. Rather than using global texture models to drive the segmentation process, we construct contourlet-based local appearance profiles at the salient points of the vertebral bodies. Non sub-sampled Contourlet sub-bands offer rich localized textural and geometric information that help localize the object of interest during the search process. A multi-scale shape prior is also constructed as described in chapter four. During the testing phase, the Contourlet-based local appearance model is used to detect vertebral boundaries in the test x-ray image. The localized multi-scale shape prior is then used to drive the segmentation process. The new method has proven to be successful in terms of its robustness to low image quality and its ability to generalize to unseen cases.

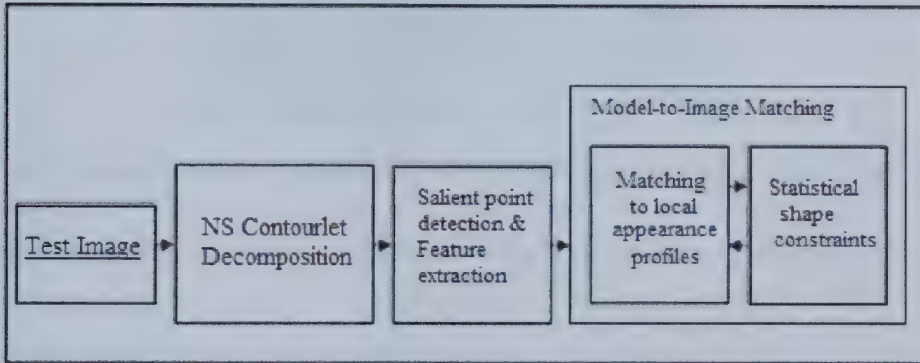
6.4.1 Methodology

The overall model learning and segmentation framework is outlined in Figure 6.8. The proposed method incorporates the following:

- i. Construction of local appearance profiles by using Non-subsampled Contourlet descriptors.
- ii. Construction of a multi-scale shape prior by using a best-basis selection framework and Independent Component Analysis (ICA).
- iii. Matching prior models to unseen images by using a shape-driven salient point matching approach.



(a)



(b)

Figure 6.8 Proposed vertebrae segmentation algorithm. (a) Model learning framework (b) Segmentation framework

6.4.1.1 Construction of local appearance profiles

Given a training set of lumbar spine x-ray images, we construct local appearance models for selected salient points that outline the vertebrae bodies. Rather than constructing the local appearance profiles for every landmark point of the shape, we build the profiles for only a few selected salient points. This approach allows for a robust and accurate model matching procedure. We refer to the selected salient points outlining the vertebral shape as the characteristic salient points (P_{ci}). Figure 6.9 shows examples of x-ray images with the outlined contour and selected characteristic salient points (seven points per vertebra).

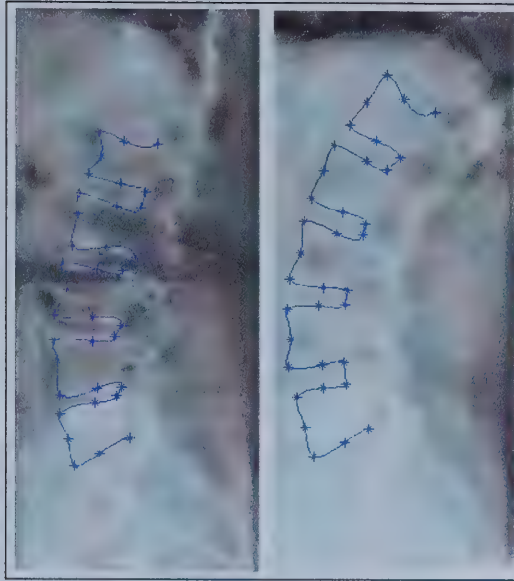


Figure 6.9 Examples of training x-ray images with outlined contours and characteristic salient points

The following summarizes the procedure to construct the local appearance profile:

i. Contourlet-based local texture descriptors

- First, every image in the training set is decomposed by using the NSCT transform. We use a 4-level NSCT decomposition. In the first level, we use a 4-level Directional Filter Bank (DFB) decomposition to get 16 directional frequency partitionings. In the second and third levels, a 3-level DFB is used to get 8 directional frequency partitionings. In the final level, 2-level DFB is used with 4 directional frequency partitionings. Figure 6.10 shows the decomposition of an x-ray image by using the NSCT transform at different scales and orientations. We refer to the response of the NSCT decompositions as the NSCT feature images.

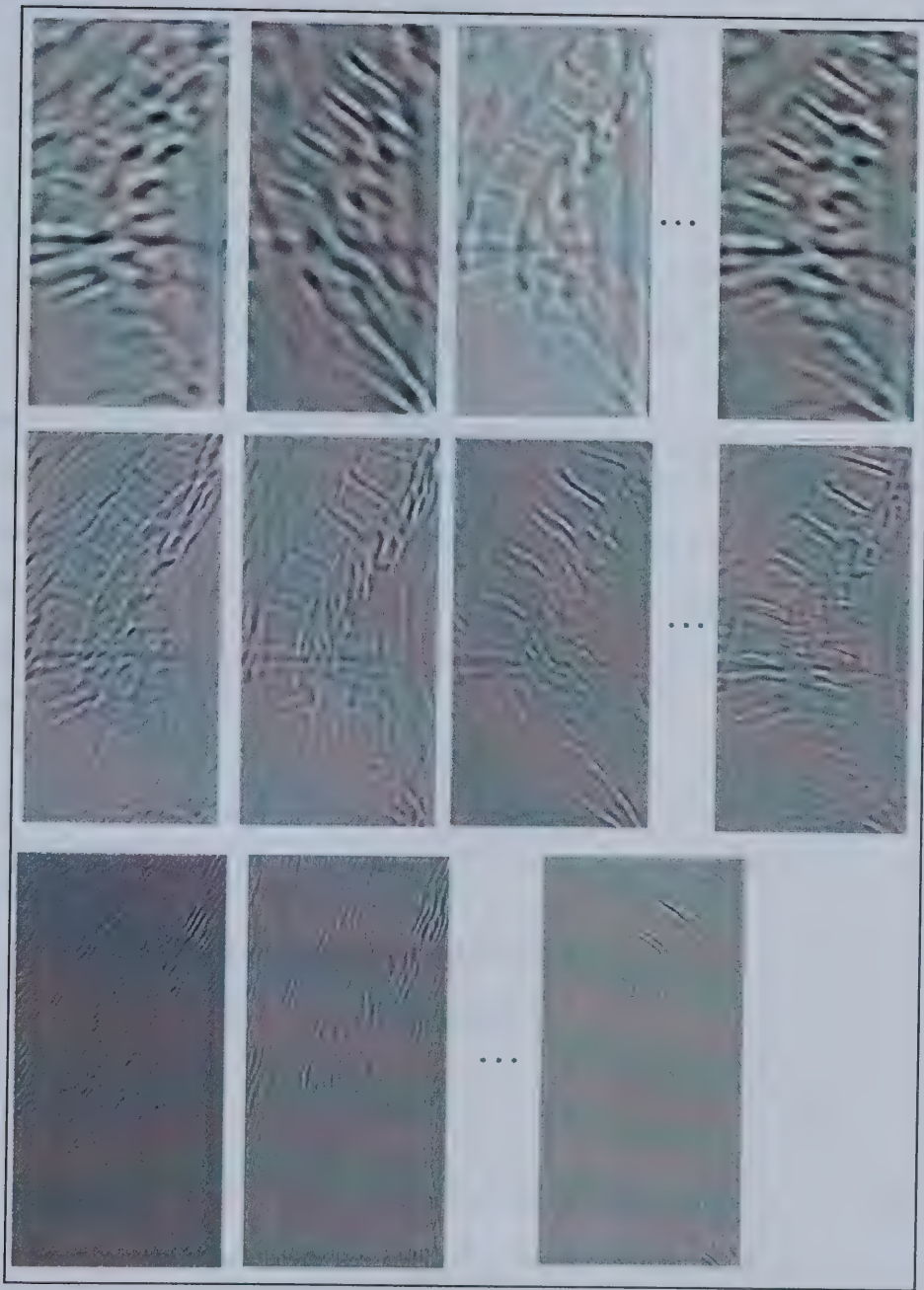


Figure 6.10 NSCT image decomposition at different scales and orientations (NSCT Feature Images).

- Second, for every characteristic salient point (P_{ci}), a descriptor of its local appearance is constructed by using the NSCT feature images. We use a square grid of $N \times N$ ($N=7$) points with the characteristic salient point at the center of the grid. A variance-based local descriptor is then extracted in the manner explained in section 6.3.2.2.

The Contourlet-based local descriptor for the characteristic salient point P_{ci} is then defined as

$$F_i = [E_1, E_2, E_3, \dots, E_D]^T, \quad (6.19)$$

where E_j is the average absolute deviation of coefficients in a square grid for the j^{th} NSCT sub-band, and D is the total number of NSCT sub-bands which is equal to 36.

ii. Profile modeling using Gaussian Mixture model (GMM)

For every characteristic salient point, the set of local descriptors extracted from the training set, $\{F_k\}_{k=1..K}$, is treated as samples in R^D . K is the size of the training set. A mixture of four Gaussian distributions is then used to represent the underlying distribution of every set of descriptors. A sample profile for characteristic point (P_{ci}) can then be written as

$$p(F_k / \theta) = \sum_{n=1}^{N_m} \alpha_n f_n(F_k), \quad (6.20)$$

where

N_m : Number of Gaussian functions in the GMM.

$f_n(F_k)$: Multivariate Gaussian density function.

α_n : Mixture weights, where $\sum_{n=1}^{N_m} \alpha_k = 1$.

The parameters of the Gaussian mixture model, $\theta = \{\alpha_n, \mu_n, \sum_n\}$, $n = 1..N_m$, are estimated from the training set by using the EM algorithm, as explained in section 6.3.2.2.

By using the training set, a local appearance model is estimated for each one of the characteristic salient points.

6.4.1.2 Construction of the multi-scale shape prior

A set of training shapes are first aligned by using the Procrustes alignment method [38].

Let the n^{th} shape vector with K landmarks be represented as

$$X_n = [x_{n1}, \dots, x_{nK}, y_{n1}, \dots, y_{nK}]^T, \quad (6.21)$$

where x and y are x - and y -coordinates.

Next, we construct the multi-scale shape prior which can capture a wide range of permissible deformations in the lumbar spine. First, we use a dictionary of wavelet-packets to find an optimal representation for the shapes in the training set. The statistics of the sparse multi-scale wavelet packet representation are then modeled by using the Independent Component Analysis (ICA). More details regarding the construction and evaluation of the multi-scale shape prior were presented in chapter four.

Any shape, X , can be described by using the multi-scale shape model as follows:

$$X = \overline{X} + (\Phi_{bb})^{-1} [\overline{\Gamma} + \Phi_{ICA} \alpha] \quad (6.22)$$

$$\Phi_{bb} = \arg \min_{\Phi \in D} (M(\Phi^T X)), \quad (6.23)$$

where

$\overline{X}$: Mean shape in the training set.

Φ_{bb} : Matrix of joint best basis for optimal representation of the shapes. Φ_{bb} is learnt from a dictionary of wavelet-packet basis by using an entropy-based cost function.

Φ_{ICA} : Matrix of statistically independent basis functions.

α : Model parameters

6.4.1.3 Segmentation and model search procedure

The local appearance profiles and the multi-scale shape prior are then used to drive the segmentation process. Figure 6.11 summarizes the model search procedure. Given an input test image, the segmentation of the lumbar spine is accomplished by using three main steps: (i) salient point detection. (ii) matching salient points to local appearance profiles, and (iii) using the multi-scale shape prior to drive the segmentation process.

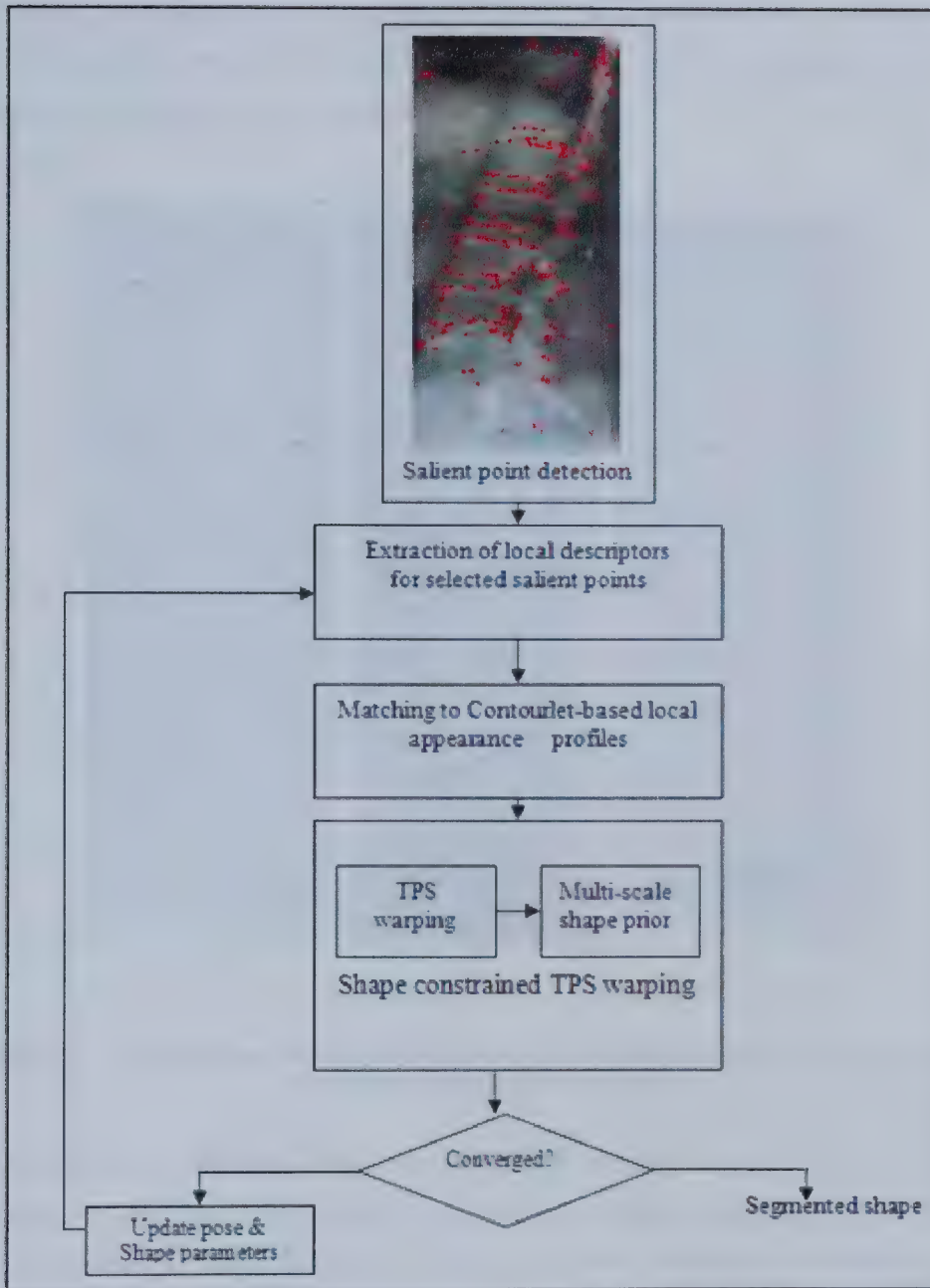


Figure 6.11 Model search procedure using Contourlet-based local appearance profiles and multi-scale shape prior.

The details are given in the following steps:

Initialization:

The search is initialized with an estimate of the approximate position of the object of interest in the image. Shape parameters are set to zero, which corresponds to the mean shape.

(a) Salient point detection: A multi-scale directional decomposition is achieved by using the Non Sub-sampled Contourlet Transform (NSCT). We use the method described in section 6.3 to extract a set of robust salient points.

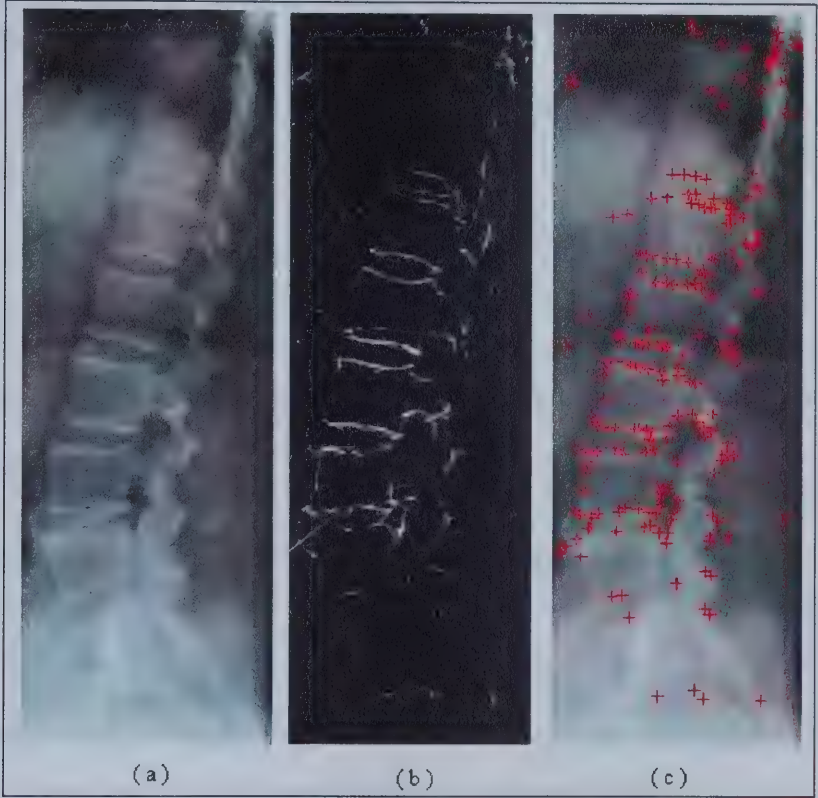


Figure 6.12 Contourlet-based salient point detection. (a) Input image, (b) strength image, (c) input image with extracted salient points.

(b) Salient point matching: Next, we use the set of local appearance profiles learnt during the training phase to locate the characteristic points outling the bodies of the vertebrae in the test image. During the search, we use the current pose estimate of the lumbar spine to restrict the search space for each characteristic point. The optimal position for any characteristic point is defined as a salient point that has a local contourlet-descriptor with the maximum probability as determined by the mixture distribution.

Each characteristic point (P_{ci}) has a local appearance profile modelled by using a Gaussian Mixture Model with the parameters $\theta_i = \{\alpha_n, \mu_n, \sum_n\}$, $n = 1..N_m$.

Let W_i denote the extent of the local search space for P_{ci} . Let $S_i = \{s_j\}_{j=1..K}$ be a subset of the salient points that belongs to the search space (W_i) of the current characteristic point.

The optimal position for the characteristic point (P_{ci}) is the salient point ($\hat{s}_j$) whose local texture descriptor (F_{s_j}) has the maximum probability as determined by the mixture distribution:

$$\hat{s}_j = \arg \max_{\{s_j\}} p(F_{s_j} / \theta_1) = \arg \max_{\{s_j\}} \sum_{n=1}^{N_m} \alpha_{jn} f_n(F_{s_j} / \mu_{jn}, \sum_{jn}). \quad (6.24)$$

(c) Shape-constrained Thin Plate Spline (TPS) warping: The outcome of step (b) is a set of selected salient points ($\{T_i\}$) that corresponds to the set of characteristic points ($\{M_i\}$, $i=1, \dots, L$) that characterizes the reference mean shape (X_M).

We use the two corresponding sets of points to estimate a smooth mapping transformation used to map the mean shape, (X_M), to the target object in the test image. By doing so, we obtain an initial segmentation of the target object.

We find a smooth transformation, f , that satisfies the following interpolation conditions at the selected salient points:

$$T_i = f(M_i), i = 1 \dots L, \quad (6.25)$$

where

M_i : is the i^{th} characteristic point on the reference shape.

T_i : is the corresponding selected salient point in the target image.

L : total number of characteristic points used to represent the object of interest.

The Thin Plate Spline (TPS) is a famous transformation that can represent smooth, rigid, and non rigid deformations [136]. Given the set of corresponding points ($\{M_i, T_i\}$),

the TPS model satisfies the interpolation condition in equation 6.25, while minimizing the bending energy ($E_{bending}$), which is the energy required to bend a thin plate spline fixed at given heights that correspond to the displacements at the points:

$$E_{bending} = \iint_{R^2} (f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2) dx dy . \quad (6.26)$$

Since the obtained corresponding point set is prone to errors, we use the approximate TPS model, which allows for an approximate interpolation between the points in favour of a smoother transformation [137]:

$$E_{TPS} = \sum_{i=1}^L [T_i - f(M_i)]^2 + \lambda E_{bending} . \quad (6.27)$$

The TPS transformation function is then given by

$$f_{TPS} = \arg \min_f E_{TPS} . \quad (6.28)$$

This minimization problem can be solved by using set of linear equations [136]. The TPS transformation is characterized by two parameter vectors $\mathbf{a} = [a_1 \ a_2 \ a_3]^T$ and $\mathbf{W} = [w_1 \ w_2 \ ...w_L]^T$. These vectors determine the affine and non-rigid part of the transformation, respectively:

$$f(x, y) = a_1 + a_2 x + a_3 y + \sum_{i=1}^L w_i U(|T_i - (x, y)|) , \quad (6.29)$$

where

$U(r) = r^2 \log(r)$ is the basis function.

$\mathbf{a}$: Parameters for the affine component of the transformation.

$\mathbf{W}$: Parameters for the non-rigid component of the transformation.

$\{T_i\}$: set of selected salient points in the target image.

Next, we obtain the segmentation estimate (X_{TPS}) by warping the mean shape (X_M) by using the obtained TPS transformation:

$$X_{TPS} = X_M \mathbf{a} + U(X_M) \mathbf{w}. \quad (6.30)$$

Figure 6.13 illustrates the estimate of the TPS warping function by using two sets of selected points. The dense mean shape is then warped by using the estimated warping function.

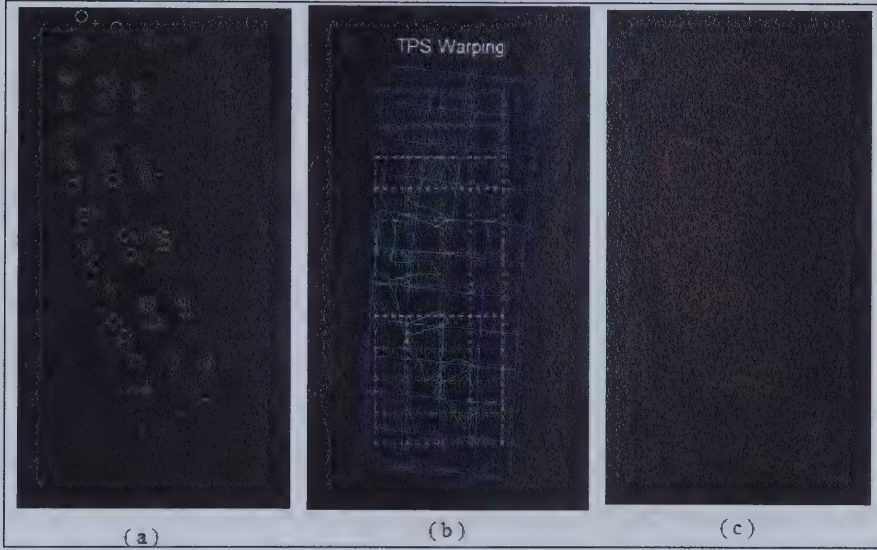


Figure 6.13 Illustration of TPS warping procedure. (a) Two sets of points used to estimate the TPS parameters, (b) estimated TPS warp, (c) Warped output shape

The TPS model tries to interpolate the selected salient points while keeping the transformation as smooth as possible. Two main factors limit the accuracy of the segmentation estimate obtained by using the TPS model alone. First, the selected salient points that the TPS model interpolates are prone to error due to the poor image quality. Second, no prior information (other than the smoothness) is imposed on the TPS transformation. Hence, it might result in non-plausible shapes.

Thus, we use the multi-scale statistical shape model to constrain the shape deformations to a space of plausible shapes. In chapter four, we illustrated the ability of the new multi-scale shape prior to capture both localized and global deformations. This ability would

enable the segmentation algorithm to generalize well in cases of normal and pathological vertebrae.

We apply the multi-scale statistical shape constraints through the following steps:

- Align the shape X_{TPS} with the mean shape with respect to rotation and scale.
- Project the shape X_{TPS} onto the space represented by the multi-scale Wavelet-ICA basis set, and obtain the model parameters $\{\hat{\alpha}\}$.
- Calculate the new statistically constrained shape estimate :

$$\hat{X}_{MS-prior} = \bar{\bar{X}} + (\varphi_{bb})^{-1} [\bar{\bar{\Gamma}} + \varphi_{ICA} \hat{\alpha}] . \quad (6.31)$$

- Apply the inverse affine transformation to $\hat{X}_{MS-prior}$ to obtain the segmentation outcome.

(d) Repeat steps (b-c) until convergence.

6.4.1.4 An alternative approach: Contourlet-based livewire best-path search

This section presents a variation of the segmentation method proposed in section 6.4.1.3. Instead of using the Thin Plane Spline (TPS) transformation to find the initial estimate of the segmented body, we propose to use a Contourlet-enhanced livewire search procedure. The livewire algorithm is a real-time segmentation method that finds the best paths between point clicks provided by the user [138]. The proposed variation integrates an enhanced livewire best-path search step in the segmentation framework described in sec 6.4.1.3. This integration is performed to benefit from the power of the salient point matching method in locating the object of interest, and also from the object delineation capabilities of the livewire algorithm. Given the set of selected salient points, the livewire method is used to find the optimal boundary that delineates the object of interest. The resulting shape is then regularized by using the multi-scale shape prior.

The livewire algorithm is a feature-based image segmentation algorithm [138]. The image is viewed as a directed graph where pixel vertices are the nodes, and the oriented pixel edges represent the edges of the graph. The problem of finding the optimal boundary between any two selected points is then equivalent to finding the minimum of a cost function between the two points [138]. The cost function is a crucial part of the livewire method. We would like the cost function to include rich information about the

boundary of the object of interest. The cost function is usually a combination of different feature images that are expected to provide information about the localization and directionality of the boundaries to be segmented. Examples of the livewire feature images in the literature include the Laplacian of Gaussian Zero crossing [138], gradient information [139], and phase information [140]. In [141], Seetharaman et al. proposed a method for detecting osteophytes in the lumbar spine by using livewire segmentation.

In this section, we propose a new Contourlet-based cost function to guide the livewire search process. Our argument is that the inherently 2D nature of the Non Subsampled Contourlet Transform (NSCT) will allow for feature images with rich geometric and texture information. Including these feature images in the cost function will enhance the ability of the livewire algorithm to localize the boundaries of the objects of interest. We use the NSCT sub-bands to extract useful local image features, which are then included in the livewire cost function. The best path search is then performed by using the new Contourlet-based cost function.

Local Image Features

- Contourlet-based edge map (E_e): The first term of the cost function is an edge map based upon the NSCT decomposition. This term is computed by using NSCT decomposition and second moment matrix analysis. This edge-map was introduced in section 6.3 to extract salient points. It is capable of capturing smooth oriented contours and corners outlining the object of interest.
- Directional Image Feature (E_{θ_p}): Directional information is also important in the livewire search algorithm. Proper directional features are needed to guide the search algorithm to take the correct local paths. Such guidance becomes important in cases involving noisy images and the presence of other attractive neighbouring structures, particularly when the human spine is involved.

The NSCT decomposes the image into a number of scales and directions. For each direction, we calculate a scale-integrated response image by first normalizing each scale level to its maximum value and forming an accumulated image by using weighted averaging:

$$E_{\theta}(x, y) = \sum_{s=1}^{N_s} \frac{1}{N_s} |R_{s,\theta}(x, y)|. \quad (6.32)$$

The outcome is a set of directional feature images capturing directional information at different orientations $\{E_{\theta}(x, y)\}, \theta = 1, \dots, N_{\theta}$. Next, we estimate the dominant orientation (θ_d), or the one with the maximum scale-integrated response. The directional image feature (E_{θ_d}) is then given by

$$E_{\theta_d} = \max_{\theta} \{E_{\theta}(x, y)\}. \quad (6.33)$$

- Edge detection (E_c): This is obtained by performing the Canny edge detector by using the Contourlet-based edge map.

Feature Combination

Before the feature images are combined, they are converted such that a low cost is assigned to the desired boundary features:

$$E = 1 - E / \max(E). \quad (6.34)$$

The new cost function is then defined as the weighted sum of all the terms described above:

$$F_{\text{cost}} = \omega_1 E_e + \omega_2 E_c + \omega_3 E_{\theta_d}. \quad (6.35)$$

Best-path search

For the second part of the algorithm, we use the Dijkstra's algorithm to find the shortest path between the selected salient points in the target image [142].

6.4.2 Comments

Section 6.4 closes with some comments regarding the proposed vertebrae segmentation method. The new segmentation method combines the power of model-based segmentation approaches with the flexibility of low-level feature-based detection methods.

- The low-level image analysis component of the proposed method is based upon the Non Subsampled Contourlet (NSCT) analysis of the x-ray image. As

discussed in chapter three, Contourlets are suited for representing higher-order singularities such as edges and contours. In medical imaging, Contourlet-based analysis provides better representation for the edges and contours of the anatomical structure of interest. Thus, Contourlet-based analysis enables us to extract robust and stable salient points for the object of interest. We also build efficient local descriptors for the selected salient points. These descriptors benefit from the rich geometrical and textural information captured by the Contourlet sub-bands.

- The proposed method employs a model-to-image fitting strategy based upon a few selected characteristic points for the object of interest. Compared to the Active Shape models and the related work in literature, our method has the advantages of increased robustness to noise. Also, the proposed segmentation method uses local texture appearance profiles to detect the object of interest. In comparison with global texture models, our model allows for more freedom in detecting target objects.
- The proposed segmentation method uses a new multi-scale shape prior. As detailed in chapter four, this prior outperforms traditional PCA-based priors. It has better generalization ability and is capable of capturing a wide range of plausible local and global deformations.

6.5 Content-based medical image retrieval using Contourlet-based salient point detector

This section presents a new method for content-based medical image retrieval in the Contourlet domain. The method integrates the new salient point detector presented in section 6.3 in a medical image retrieval framework.

6.5.1 Background

Over the past decade, researchers have been increasingly interested in content-based medical image retrieval (CBIR) methods [143]. CBIR systems aim at retrieving images from a database by using visual image features rather than textual information. With the growing number of digital medical images in hospitals, efficient CBIR techniques can benefit a wide range of medical practices. However, medical image retrieval is a challenging task due to the quality of medical images, the complexity of anatomical structures, and the wide range of intra-class variability.

Traditionally, various global and local features have been used for image retrieval applications. Examples of global features include color information, histograms, and response from directional filters such as Gabor filters. In the field of medical imaging, local features appear to be more appealing than global features. Moreover, using local features in medical image retrieval opens the door for pathological-based retrieval systems. In [144], Deseleers et al. presented a comprehensive review of the features employed in image retrieval systems. In [87], Antani et al. investigated using different shape features of the human vertebrae for the content-based image retrieval of spine x-ray images. In [145], Liu and Tong proposed a method for medical image retrieval using Gabor-based textural features extracted from salient points. In [146], Maree et al. proposed a method for image retrieval using randomly selected sub-windows. In [147], Tommaasi et al. suggested using SIFT features for medical retrieval applications. In [148], Howarth suggested using Gabor-based texture features and color information for medical image retrieval. In [149], Ngo suggested using DCT-based features. In [150], Selvarani et al. proposed a method for medical image retrieval using Gabor-based low-level features and high-level semantic features.

6.5.2 Methodology

This section presents the details for the new content-based medical image retrieval system. We propose a method for content-based medical image retrieval using a NSCT-based salient detector and local descriptors. The NSCT-based salient point detector is used to extract salient points that locally characterize the object of interest. We use local descriptors to capture local textural information from the multi-scale multi-directional sub-bands of the Non Subsampled Contourlet domain. Figure 6.14 summarizes the proposed medical image retrieval system.

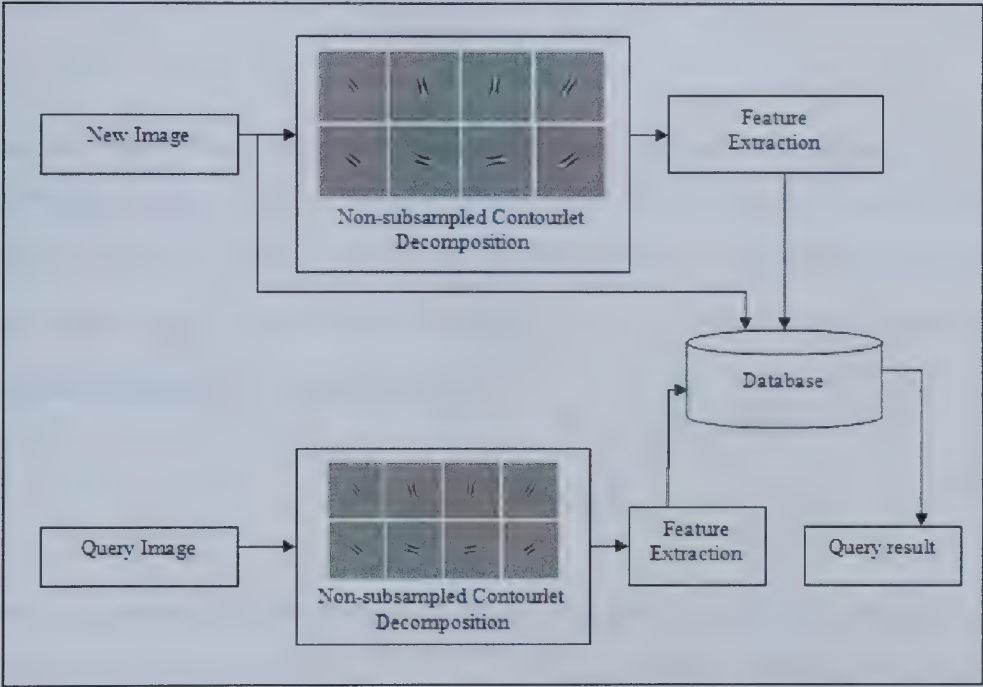


Figure 6.14 Content-based Image retrieval using Non-subsampled contourlet features

The following outlines the details of the proposed method:

- We perform multi-scale multi-directional image decomposition by using a Non-subsampled Contourlet Transform (NSCT). We use 3 decomposition scales with 4-, 3-, and 2- directional levels. This corresponds to a total of 28 sub-bands.

- Next, we apply the salient point detector proposed in section 6.3 to extract robust and stable salient points. We extract a set of N salient points, $\{SP_i\}, i=1,2,...N$, from each image in the database. In our experiments, N was set to 100.

- For each salient point (SP_i), we extract a local descriptor from the NSCT sub-bands.

Let V_{ij} represent the local descriptor extracted from the i^{th} salient point in the j^{th} image. Then, the feature vector (F_j) that represents the j^{th} image is formed by concatenating the extracted local descriptors from all N salient points:

$$F_j = [V_{1j} \ V_{2j} \dots V_{Nj}]^T. \quad (6.36)$$

- The final feature vector is generated through a dimension reduction step using Principal Component Analysis (PCA). Using a training set of M images, we extract the feature vectors, $\{F_j, j=1,2,...M\}$. Then, we use the feature matrix to estimate the eigenvector basis matrix (Φ_{PCA}). Any NSCT-based feature vector (F_j) can be represented by the projections onto the PCA-basis matrix Φ_{PCA} :

$$\hat{F} = \overline{F} + \Phi_{PCA} \alpha, \quad (6.37)$$

where α represents the contributions of various Principal Components to the variance in the training dataset. Dimension reduction is then achieved by retaining only the PCA coefficients that capture up to 90 % of the variance in the dataset. Given the PCA basis matrix (Φ_{PCA}) and the original feature vector of a new image (F_{new}), the final feature vector is obtained by projecting onto the PCA basis functions that retain only 90 % of the variance:

$$\alpha_{new} = \Phi_{PCA}^T (F_{new} - \overline{F}). \quad (6.38)$$

- The final step is the image retrieval. Given a query image, we use the Euclidean distance to measure the similarity between features extracted from the query image, and others in the database. The obtained distances are then sorted to obtain a ranked list of retrieved images:

$$ED_q = \sqrt{(F_q - F_j)^2}, \quad (6.39)$$

Where ED_q is the Euclidean distance between the query image I_q and the j^{th} image in the database.

6.5.3 Comments

This section highlights some of the important aspects of the proposed content-based image retrieval system.

- The proposed framework is based upon local features extracted from the Non-Subsampled Contourlet domain. The NSCT-based image analysis allows for the extraction of rich textural and geometric information.
- Using the NSCT-based salient point detector at the core of the retrieval system allowed for the extraction of robust and repeatable salient points suitable for medical images.
- Due to the locality of the features extracted, the system has the potential to allow for more sophisticated retrieval tasks such as pathology-related retrieval.

6.6 Experiments and Results

In this section, we evaluate the performance of the methods presented in this chapter. In particular, experiments were conducted to evaluate the following:

- Robustness of the NSCT-based salient point detector and local descriptor.
- Accuracy of the new vertebral segmentation algorithm.
- Accuracy of the new medical image retrieval method.

6.6.1 Evaluation of the NSCT-based salient point detector

Datasets:

We evaluate the performance of the proposed NSCT-based salient point detector by using standard test images. Figure 6.15 shows examples of the images used in the experiments. Figure 6.15(a) shows a standard synthetic image that exhibits different boundaries. Figure 6.15(b) is a VanGogh image, which is a famous image with many texture variations. Figure 6.15(c) shows an example of the lumbar spine x-ray images. These images are used to quantify the performance of the new detector in terms of the following:

- Localization capabilities.
- Robustness to geometric and photographic transformations.
- Robustness to different noise levels.
- Efficiency in medical image analysis.

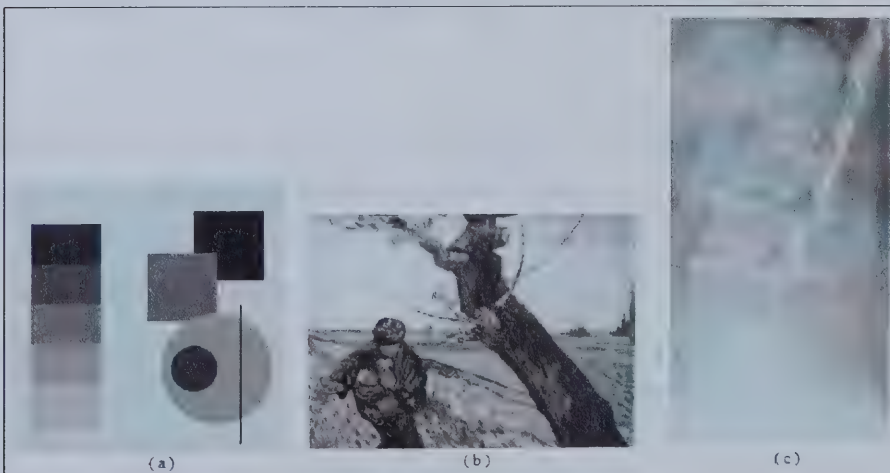


Figure 6.15 Images used for evaluation of the NSCT-based salient point detector. (a) Synthetic image, (b) VanGogh Image, (c) X-ray image of lumbar spine

Evaluation Criterion: Repeatability Rate

We evaluate the robustness of the proposed salient point detector by using the Repeatability Rate criterion (rr) introduced by Schmid et al. [126].

Given two images related together by some transformation, the repeatability rate is the ratio between the number of corresponding points detected in both images to the total number of points in the common area. In our experiments, we consider the following relationships between the reference and test images:

- Rotation transformation.
- Thin Plate Spline transformation (as an example of non-rigid transformation).
- Variations in illumination levels.
- Variations in noise levels.

Let $\phi^{(r)} = \{X_i^{(r)}\}$ and $\phi^{(t)} = \{X_j^{(t)}\}$ be the two sets of interest points detected in the reference and test images, respectively. Let N_c be the number of salient points in the common area. Let $C\{X_i^{(r)}, X_j^{(t)}\}$ be the number of corresponding pairs of salient points with a localization error of 1.5 pixels. The Repeatability rate (rr) between the two images is defined by

$$rr = \frac{C\{X_i^{(r)}, X_j^{(t)}\}}{N_c}. \quad (6.40)$$

We evaluate the proposed salient point detection method (*NSCT-based*) against the following famous salient point detectors from the literature:

1. The Harris detector (*Harris*).
2. A detector proposed by Kovesi et al., based upon Log-Gabor decomposition and phase congruency (*PC-based*) [134].
3. A variation of the proposed salient point detector based upon steerable pyramid decomposition (*Steerable-based*).

6.6.1.1 Localization

The localization capability of the new salient point detector is evaluated by using the synthetic test image shown in Figure 6.15(a). Figure 6.16 shows the salient points extracted by using various detectors. The proposed NSCT-based detector is capable of extracting most of the salient points in the image. The Harris detector and the Laplacian-of-Gaussian (LOG) detector seem to perform well only in case of strong salient regions (corners and edges). However, the two detectors show a limited localization capability in the cases involving weak edges and smooth contours. The phase congruency method (PC-based) demonstrates a better performance compared to that of the Harris and LOG detectors. The proposed NSCT-based detector is more capable of localizing the salient points along the smooth circular contour. This superior capability is attributed to the fact that the Non-subsampled Contourlet Transform (NSCT) permits a more flexible directional analysis of the image. Similar observations were obtained when the detectors were applied to a spine x-ray image. Among all the detectors investigated, the proposed detector seems to best detect anatomical structures with poor edges and smooth contours.

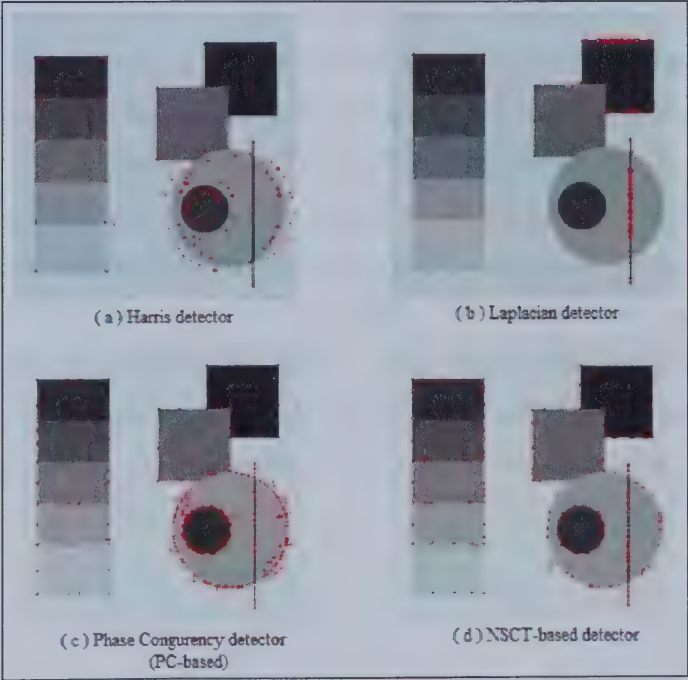


Figure 6.16 Localization of various salient point detectors. (a) Harris detector, (b) Laplacian detector, (c) Phase Congruency detector (PC-based), and (d) NSCT-based detector.

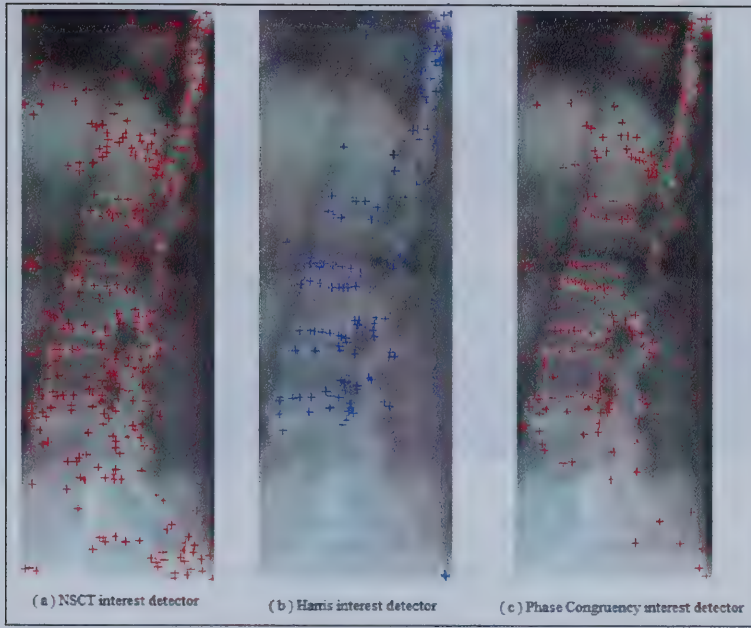


Figure 6.17 Localization of various salient point detectors in medical images. (a) NSCT-based detector, (b) Harris detector, and (c) Phase Congruency detector.

6.6.1.2 Robustness to rotation

This section evaluates the performance of the proposed salient point detector in terms of its robustness to the rotation angle. Experiments were conducted by using a set of rotated VanGogh images. Rotations are obtained by rotating the image around its central axis. The image set can be obtained from [151]. The repeatability rate is calculated as a function of the rotating angle over the range of 10-180 degrees. Figure 6.19 shows the performance of the proposed interest point detector in comparison to a number of famous interest point detectors from the literature [126]. The Harris detector and the Cottier detector are based upon the auto-correlation matrix. The Heitger detector is based upon Gabor filters. In comparison to the other methods, the proposed NSCT-based detector has achieved the highest rates of repeatability over the entire range of angles.



Figure 6.18 VanGogh test image at various rotation angles.

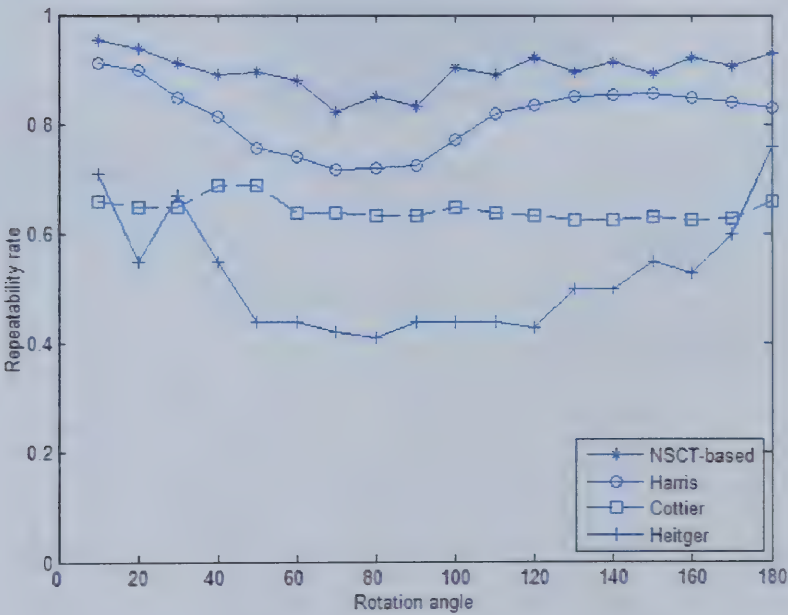


Figure 6.19 Robustness to rotation of various salient point detectors

6.6.1.3 Robustness to noise

The robustness of the proposed salient point detector to noise is evaluated by using VanGogh test images with added Gaussian noise. Figure 6.20 shows examples of noisy VanGogh test images with different noise levels. Figure 6.21 shows the effect of variations in noise levels on the repeatability rate for the proposed NSCT-based detector and the Harris detector. At noise variance levels below 0.02, the new method has achieved an average improvement of 15.23 % over the Harris detector. The proposed

method demonstrates much more robustness at increased noise levels. At noise variance levels above 0.02, an average improvement in the repeatability rate of 70.40 % has been achieved.

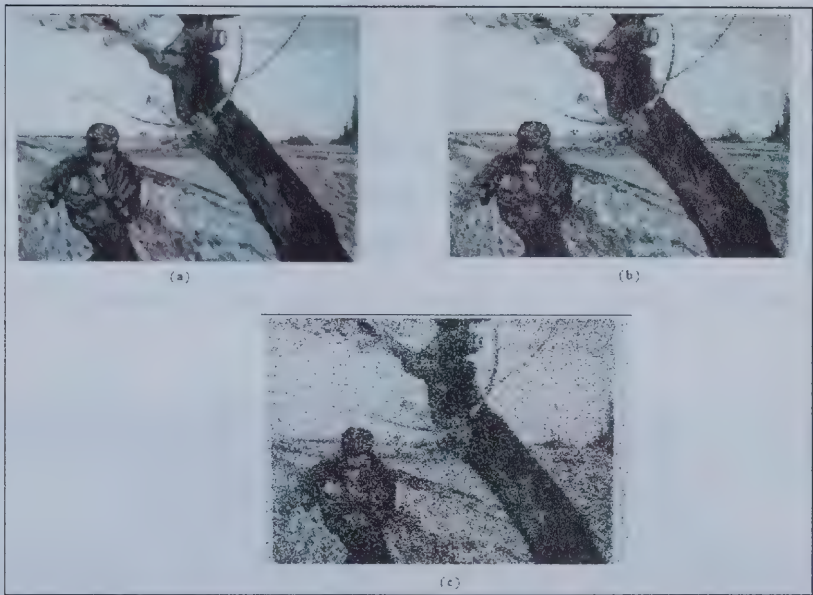


Figure 6.20 VanGogh test image with added Gaussian noise : (a) clean image, (b) 0.01 noise variance, (c) 0.06 noise variance

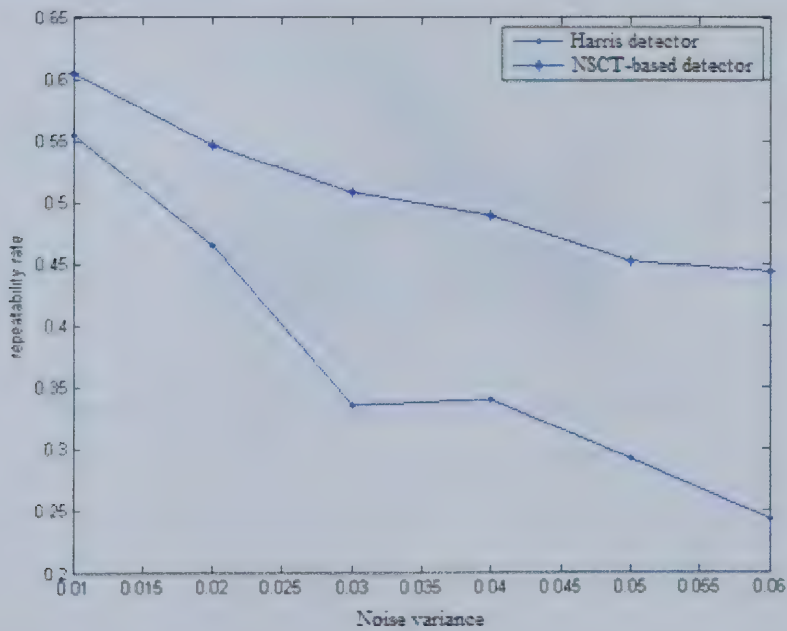


Figure 6.21 Effect of variations in noise levels on the repeatability rate for two different salient point detectors

6.6.1.4 Robustness to change in illumination

The robustness of the proposed method to a uniform change in illumination is evaluated by using the VanGogh test image with changing levels of brightness. The brightness levels are changed by changing the illumination factor (β). Figure 6.22 shows the VanGogh image with different brightness levels. A positive value of β increases the brightness level, and a negative value of β decreases the brightness level. Figure 6.23 shows how the repeatability rate is affected by altering the brightness level of an image. Compared to the Harris detector, the proposed NSCT-based detector achieves an average improvement in repeatability rate of 48.89%. Our method achieves an improvement of 13.124 % in the repeatability rate over the phase congruency detector (PC-based) [134].

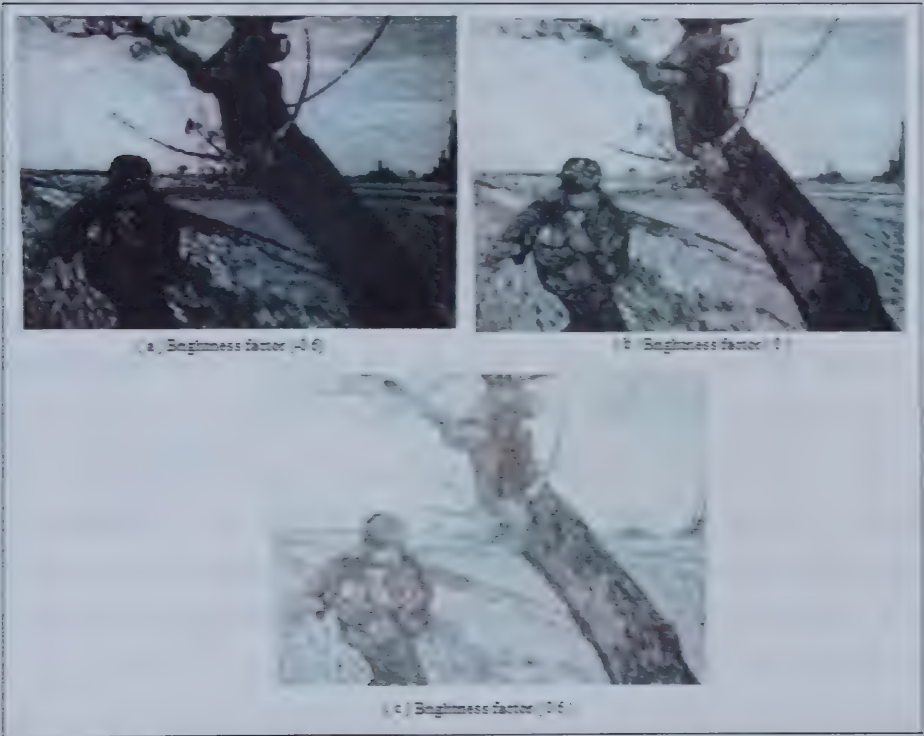


Figure 6.22 VanGogh test image with various levels of brightness

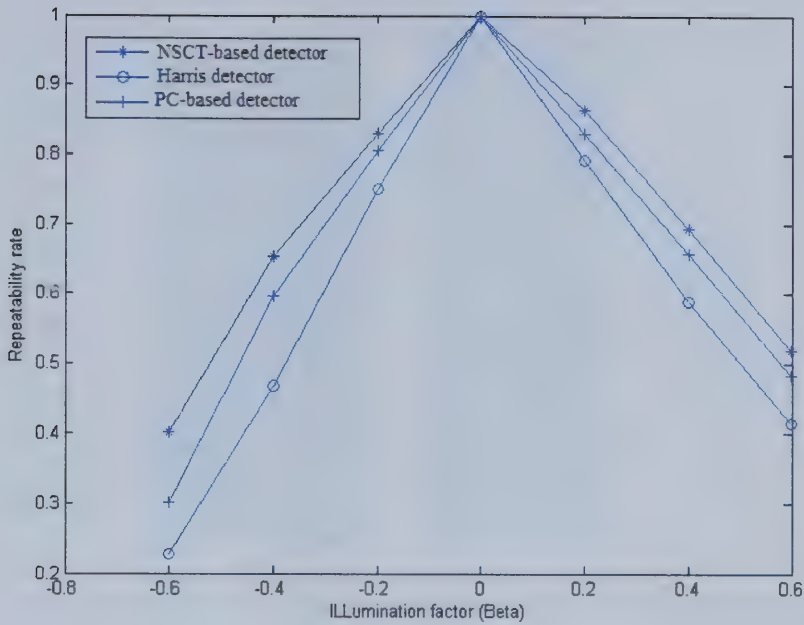


Figure 6.23 Effect of brightness level on the repeatability rate for various salient point detectors

6.6.1.5 Robustness to non-rigid transformations

In medical images, anatomical structures undergo a wide range of non-rigid deformations. Hence, a good salient point detector for medical images needs to be robust to non-rigid deformations in the object of interest. Thus, we evaluate the robustness of the proposed salient point detector to non-rigid deformations in the lumbar spine. Experiments are conducted by using a set of x-ray images of the lumbar spine with different non-rigid deformations. The object of interest in every image pair is related by a non-rigid transform. The Thin Plate Spline model is used to estimate the non-rigid transformation. We calculate the repeatability rate as the percentage of correct matching with respect to the total number of common points between the two images. Table 6.1 shows the resulting average repeatability rates obtained by using the proposed NSCT-based detector, Harris detector, PC-based detector, and the steerable-based detector.

The proposed NSCT-based salient point detector achieves improvements of 64.22 % and 18.01 % in the average repeatability rate in comparison to the Harris detector and the PC-based detector, respectively. The NSCT-based detector achieves an improvement of 9.95% in comparison to the steerable-based detector.



Figure 6.24 Two x-ray images where the object of interest is related by non-rigid deformations

Table 6.1 Evaluation of robustness of different salient point detectors to non rigid deformations of the object of interest.

Salient Point detector method	Average repeatability rate
NSCT_based	0.7891
Harris detector	0.4805
PC-based detector	0.6687
Steerable-based	0.7177

6.6.1.6 Appearance-based salient point detector

In section 6.3.2.2, we presented a method to incorporate prior local appearance knowledge of the salient points in the detection process. We evaluate the efficiency of the proposed method in terms of the average reduction ratio in the number of detected salient points. We compare the performance of the proposed NSCT-based salient point detector in two cases: with prior appearance knowledge and without prior appearance knowledge. We use both detectors to extract salient points from x-ray images of the lumbar spine. The targets of interest in the images are the bodies of the lumbar vertebrae. The use of the proposed appearance-based detector results in a significant reduction in the number of salient points by excluding salient points not relevant to the object of the interest. Figure 6.25 shows examples of salient points extracted from an x-ray image by using the two methods. An average reduction ratio of 51.86 % is achieved by using the appearance-based salient point detector.

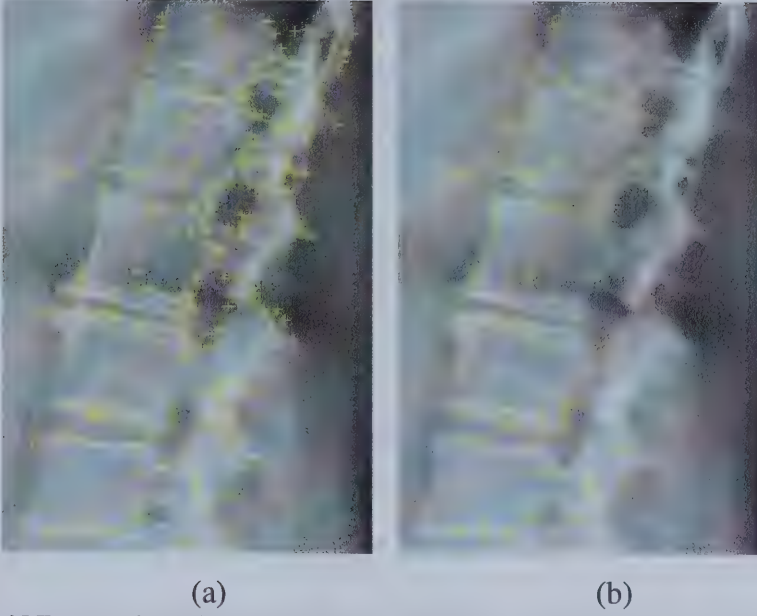


Figure 6.25 Example for salient point detection using NSCT-based method (a) with no prior appearance knowledge, (b) with prior appearance knowledge.

6.6.2 Evaluation of contourlet-based local descriptor

The proposed contourlet-based local descriptor appears at the core of the methods for vertebrae segmentation and image retrieval. The efficiency of the local descriptor greatly affects other higher level tasks. In this section, we evaluate the performance of the proposed local descriptor and assess its ability to capture localized textural and geometrical information.

Evaluation criteria

We use the “Matching score” as the evaluation criterion. Given two images that are related together by some transformation, the matching score is defined as the ratio between the number of correct matches to the total number of matches.

Let $\phi^{(r)} = \{X_i^{(r)}\}$, and $\phi^{(t)} = \{X_j^{(t)}\}$ be two sets of points in the reference and target images, respectively. Each point (X_i) is described by using the NSCT-based local descriptor (F_i). Thus, we have two sets of local descriptors ($\Lambda^{(r)} = \{F_i^{(r)}\}$, $\Lambda^{(t)} = \{F_j^{(t)}\}$) that correspond to sets of points in the reference and the target images, respectively. A point in the reference image ($X_i^{(r)}$) is

said to match a point in the target image ($X_j^{(r)}$) if the Euclidean distance between the two corresponding local descriptors is less than a threshold:

$$\left\| F_i^{(r)} - F_j^{(t)} \right\|_2 < Threshold . \tag{6.42}$$

Any match found is said to be correct if the two points (X_i and X_j) spatially correspond to each other within a localization error of 1.5 pixels.

We compare the performance of the proposed NSCT-based local descriptor against two relevant descriptors from the literature:

1. GaborJets descriptor [145]: This is a local descriptor constructed by using Gabor filters.
2. Steerable-based descriptor [152,153]: A local descriptor based upon steerable pyramid decomposition.

6.6.2.1 Robustness to noise

We use the VanGogh test image with added levels of Gaussian noise to assess the performance of the proposed NSCT-based descriptor in the presence of noise. We calculate the matching scores among a number of VanGogh images that have different levels of noise. Table 6.2 shows the average matching score for the proposed descriptor, the steerable-based descriptor, and GaborJets. An average matching score of 96.05 % is achieved by using the proposed descriptor, versus an average matching score of 86.43 % and 87.74 % for the steerable-based and GaborJet descriptors, respectively.

Table 6.2 Robustness of various local descriptors to noise

Local Descriptor	Average Matching score
NSCT-based	96.05 %
Steerable-based	86.43 %
GaborJets	87.74 %

6.6.2.2 Robustness to change of view

We also evaluate the performance of the NSCT-based descriptor against a change in the viewing angle. We use a set of standard test images with varying viewing angles [151]. Figure 6.27 shows the matching score as a function of the change in viewing angle. On general, the matching score tends to decrease with the increase of the difference between

the viewing angles. Among the local descriptors investigated, the proposed NSCT-based descriptor seems to be least affected by a change in the viewing angle.

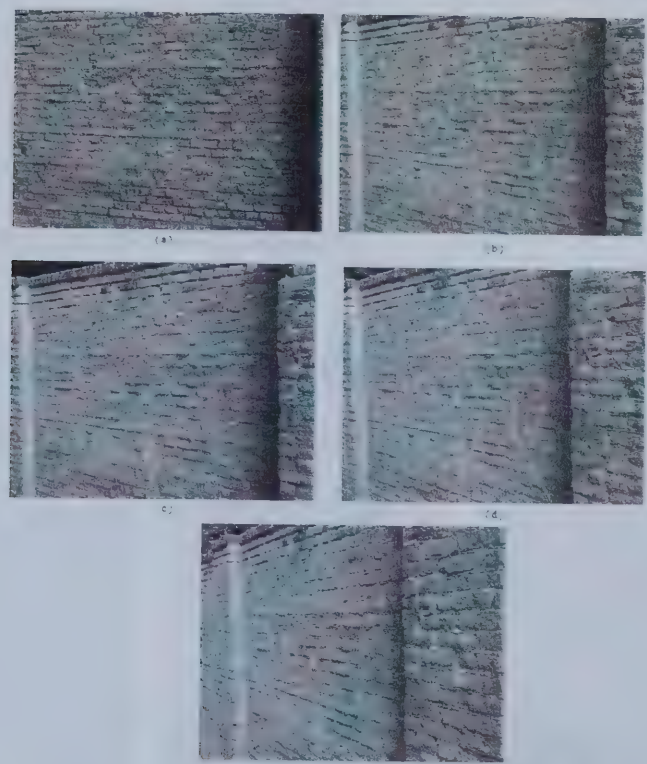


Figure 6.26 Test images with changing viewing angle

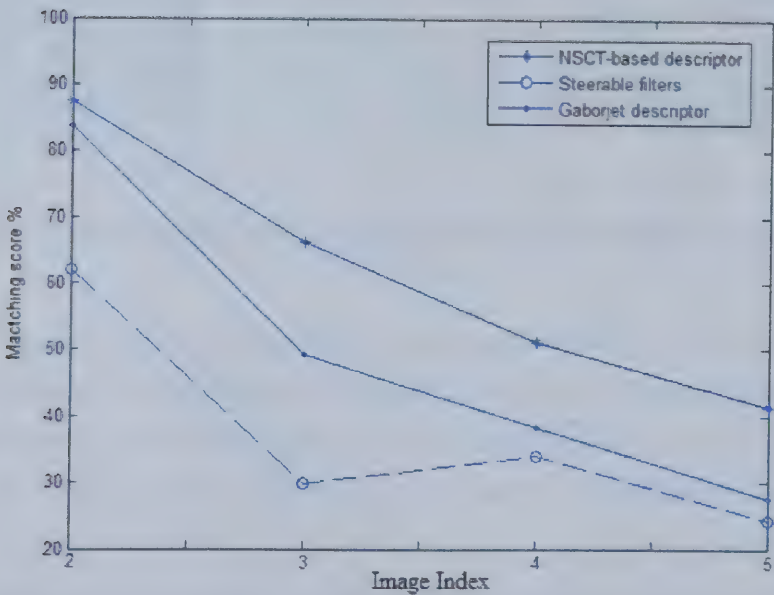


Figure 6.27 Robustness of various salient point detectors to change in viewing angle

6.6.3 Evaluation of segmentation algorithm

The vertebrae segmentation experiments were performed by using a dataset of 100 x-ray images of a lumbar spine with outlined contours of the five lumbar vertebrae L1-L5. The x-ray images were provided by the National Library of Medicine [108]. Our results are presented in this section. The accuracy of the proposed vertebrae segmentation method is assessed by using a leave-40-out experiment, where 60 images are used in the training phase, and 40 images are used for testing. The segmentation accuracy is quantitatively evaluated in terms of the point-to-contour error.



Figure 6.28 Examples of x-ray images of the lumbar spine used in evaluation of the segmentation algorithm

Figure 6.29 compares the proposed segmentation method to the Active shape model (ASM). The first row shows examples where the ASM fails to accurately capture various deformations in the lumbar spine. The segmentation results in the second row reveal that the proposed method is able to segment a wider range of deformed vertebrae. This is attributed to two main reasons: (i) the effectiveness of the proposed Contourlet-based model-to-image fitting procedure, and (ii) the generalization ability of the multi-scale shape prior.

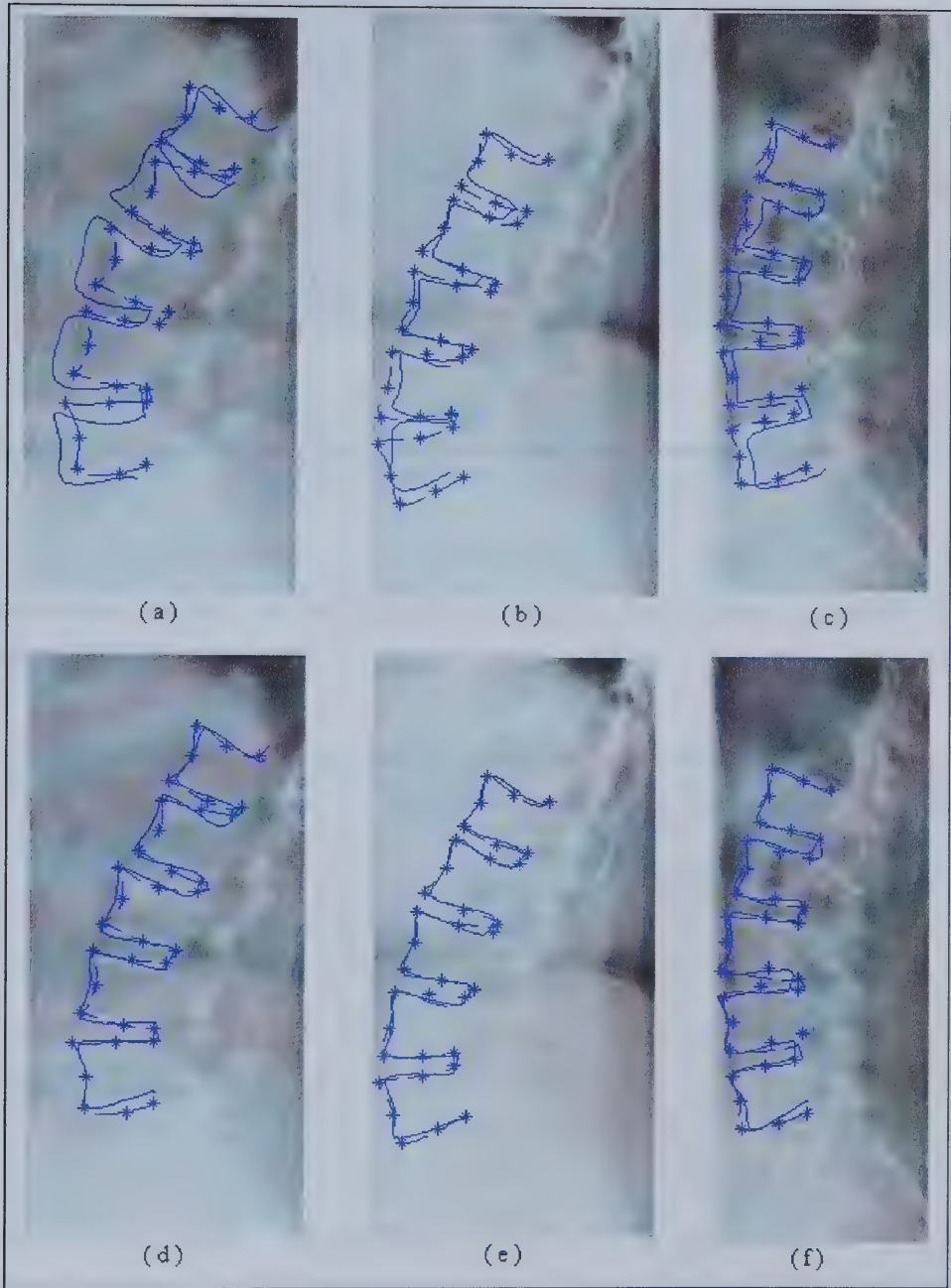


Figure 6.29 Segmentation results for three different images by using Active Shape model (ASM) (a,b,c), and the proposed segmentation method (d,e,f).

Figure 6.30 illustrates the effect of using the new multi-scale prior in the proposed segmentation framework. This figure reveals that the new multi-scale shape prior is capable of capturing local deformations in the vertebral bodies.

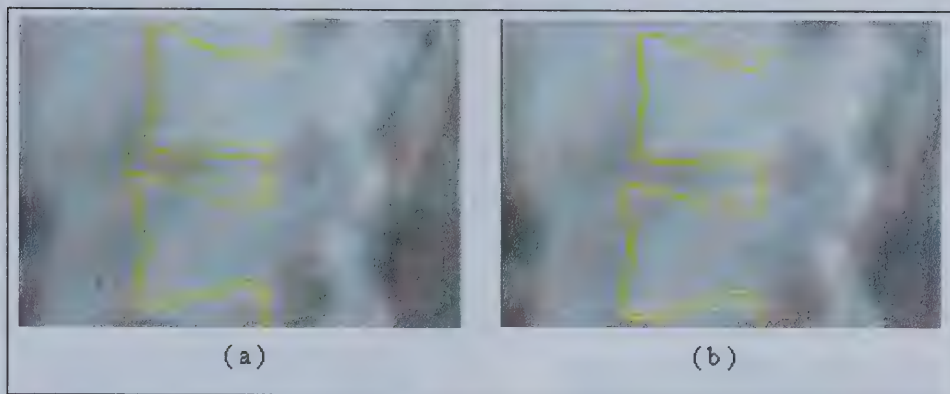


Figure 6.30 Effect of using the proposed Multi-scale prior on segmentation accuracy (a) No prior is used, (b) localized multiscale prior is employed

Table 6.3 summarizes the segmentation accuracy per vertebra for the proposed segmentation method. The combined average segmentation accuracy for all the lumbar vertebrae is 2.669 mm.

Table 6.3 Segmentation accuracy per vertebra for the proposed method

Vertebra	Average pt-to-contour error	Max pt-to-contour error
L1	1.3270 mm	2.7654 mm
L2	1.1510 mm	2.2063 mm
L3	1.2630 mm	2.1028 mm
L4	1.2303 mm	2.2487 mm
L5	1.3898 mm	4.0208 mm
Combined	1.2722 mm	2.669 mm

We also investigate the potential gains from using the proposed Contourlet-based livewire procedure to delineate the vertebral bodies. Figure 6.31 compares the delineation capabilities of the proposed Contourlet-based livewire method with those of the traditional livewire algorithm. The comparison reveals that the contourlet-based livewire exhibits better delineation capabilities, especially for weak edges and contours.

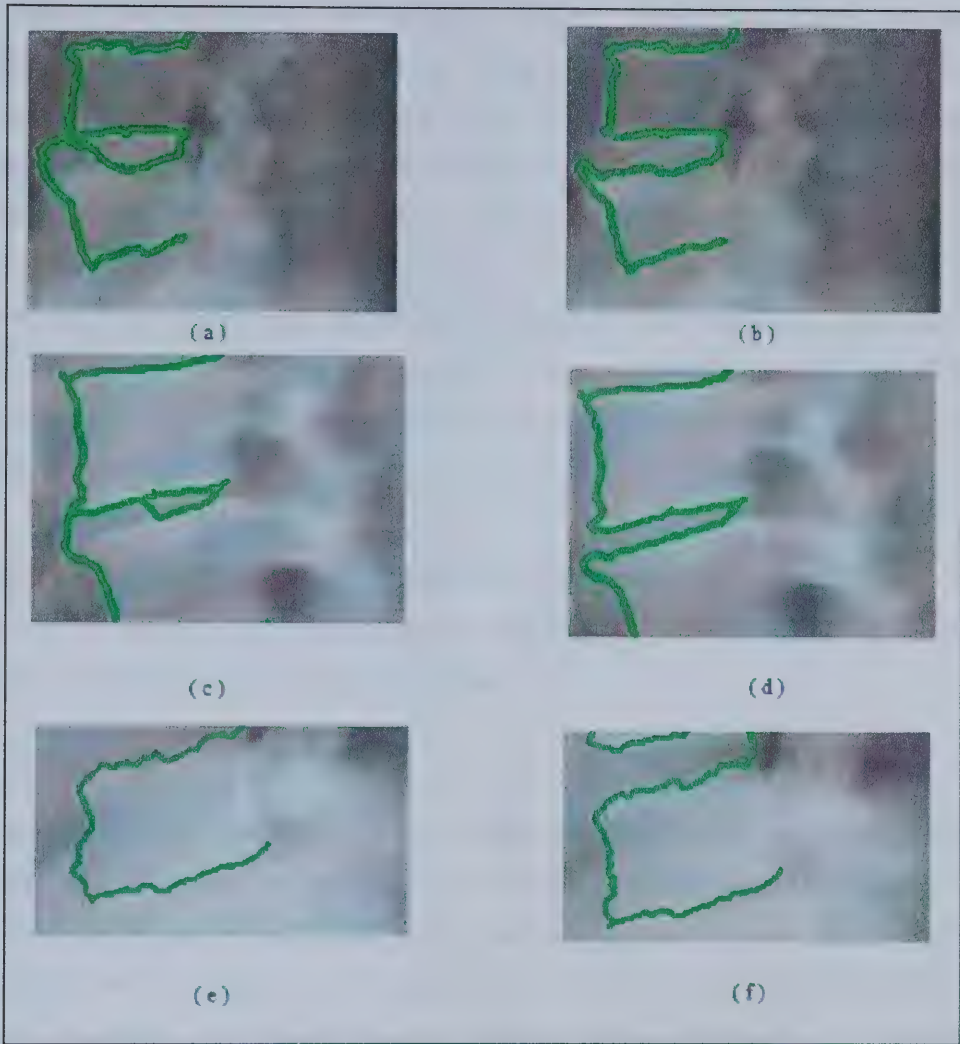


Figure 6.31 Comparison for the delineation capabilities of the proposed contourlet-livewire method (b, d, f), and the traditional livewire algorithm (a, c, e).

Discussion

We presented a novel method for vertebrae segmentation using a shape-driven salient point matching procedure. The proposed method incorporates the following: (i) a novel Contourlet-based local appearance model used to detect vertebral boundaries in the test x-ray image, and (ii) a localized multi-scale shape prior used to drive the segmentation process. An average accuracy of 2.669 mm is achieved.

Our results compare well with those of other vertebrae segmentation methods in the literature [36, 74, 56, 82, and 154]. In [154], Zomaro et al. proposed a hierarchal segmentation with an accuracy of 3mm for cervical vertebrae, and 6.4 mm for lumbar vertebrae. This method is fully automated since the Generalized Hough transform is used

to estimate the initial pose of the spine. In [56], Roberts and Cootes reported a segmentation accuracy of 1.28 mm by using a method based upon linking Active Appearance sub models. This method involves building multiple appearance models for segments of the spine to make up for the lack of localization in traditional appearance models. In [74], De Bruijne et al. reported a mean point-to-contour error of 1.4 mm for lumbar spine segmentation using shape particle filtering. This method involves a time-consuming pixel classification stage. Thus, it is more suited for offline applications. In [82], Howe et al. reported a segmentation accuracy of 4.35 mm for lumbar spine segmentation using the Active Appearance model and the Generalized Hough transform to estimate the initial pose.

6.6.4 Evaluation of the proposed medical image retrieval method

This section presents our experiments to evaluate the performance of the proposed content-based medical image retrieval system.

Dataset:

We use images from the IRMA database (ImageClef2007) [125], which is publically available as a part of the Image Retrieval for Medical Applications (IRMA) project [125]. The database consists of radiograph images randomly collected from regular routines in the department of diagnostic Radiology, Aachen University of Technology (RWTH), Aachen, Germany. The images represent various views, pathologies, and anatomical structures. The complete database consists of 11,000 images with ground truth hierarchal classification. In our experiments, we use a set of 1000 images. 800 images were used in training, and 200 images were used as query test images. Figure 6.32 shows examples of the images used in our experiment.



Figure 6.32 Examples of the IRMA medical image database (Courtesy of TM Deserno, Dept. of Medical Information, RWTH, Aachen, Germany [125])

Evaluation Criteria

We evaluate the performance of the new method by using the following criteria:

- *Average retrieval rate (ARR)*: For a query image, the retrieval rate is the ratio between the number of relevant images retrieved to the total number of retrieved images. The average retrieval rate (*ARR*) is calculated by averaging the retrieval rates over all the query images:

$$RR_q = \frac{N_{\text{relevant-retrieved}}}{N_{\text{retrieved}}} \quad (6.43)$$

$$ARR = \sum_{q=1}^{N_q} RR_q, \quad (6.44)$$

where $N_{\text{relevant-retrieved}}$: Number of relevant images retrieved,

$N_{\text{retrieved}}$: Total number of retrieved images,

RR_q : Retrieval rate for the q^{th} image.

- *Average Error Rate (ER)*: This is the average percentage of first retrieved images that are not relevant images. This rate is equivalent to the misclassification rate of a nearest neighbour classifier using the same distance measure.

Results:

Figure 6.33 shows examples of images of the same class where localized retrieval methods are more efficient than global retrieval methods. Despite being from the same class, the three images exhibit a wide range of variations in quality, pose and deformations.

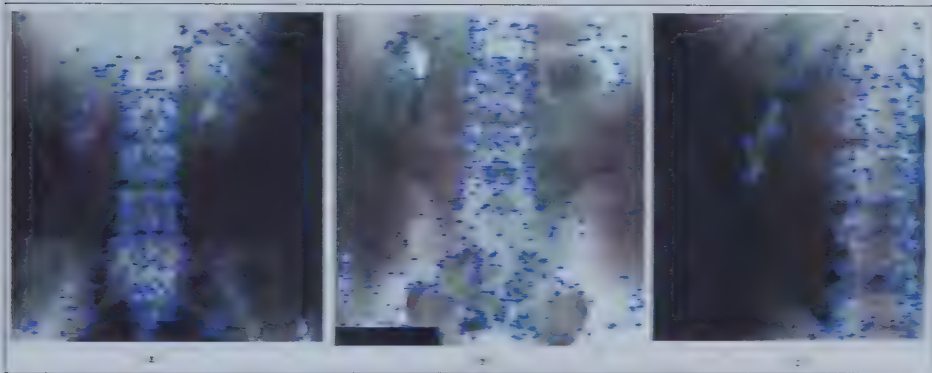


Figure 6.33 Example of intra-class variations for images in the IRMA database

Figure 6.34 evaluates the performance of the proposed CBIR system by using average retrieval rates (ARR) as a function of the number of top retrieved images. We compare the performance of the following systems:

- i. The proposed retrieval system using Contourlet-based salient point detectors (NSCT-local).
- ii. A retrieval system based upon Gabor-based salient detectors (Gabor-local).
- iii. A retrieval system based upon the global features extracted from the response of the NSCT sub-bands (NSCT-global).

The proposed retrieval method (NSCT-local) tends to consistently outperform the other two methods. Over a range of 5–40 top retrieved images, the proposed method achieves an average improvement in the average retrieval rate (ARR) of 9.574 % and 17.338 % over the Gabor-based system and the NSCT-global system, respectively.

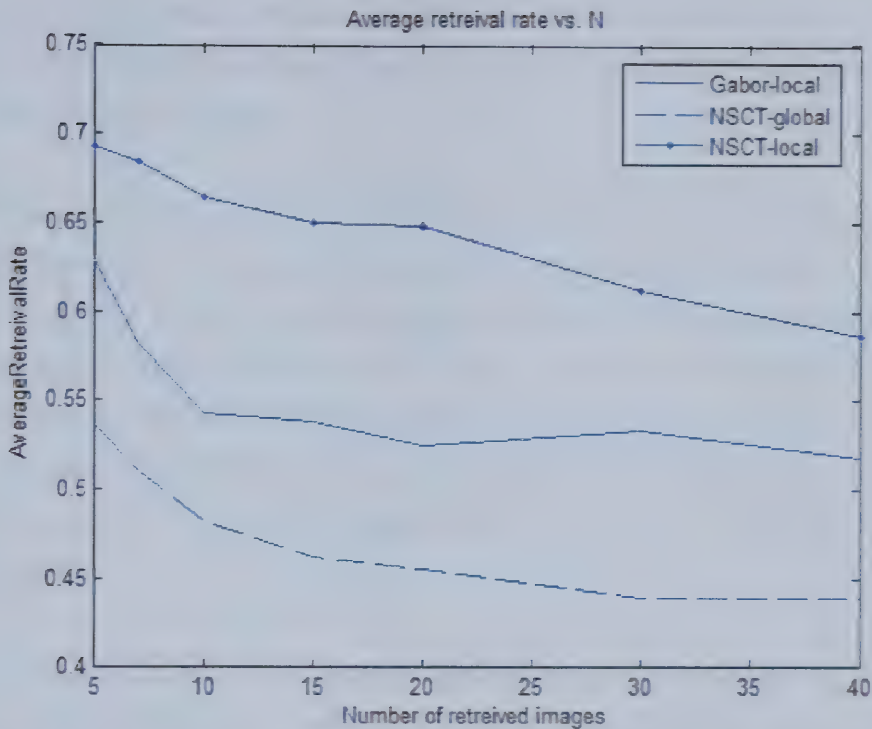


Figure 6.34 Average retrieval rates versus number of retrieved images for the following CBIR methods: Gabor-based salient points (Gabor-local), NSCT-based salient points (NSCT-local), and NSCT-based global response (NSCT-global).

Table 6.4 evaluates the performance of the proposed method in terms of the average error rate (AERR). The proposed method (NSCT-local) achieves an average error rate of 21.43 % versus 35.71 % and 46.43 % for the Gabor-based system, and the NSCT-global system, respectively.

Table 6.4 Average error rates for different CBIR methods

Method	Average Error Rate
NSCT-local	21.43 %
Gabor-local	35.71 %
NSCT-global	46.43 %

Discussion

In this section, we evaluated the performance of the new Content-based medical image retrieval system. The new system is based upon local features extracted from the Non-subsampled Contourlet domain. Using the NSCT-based salient point detector as the basis of our retrieval system allowed for the extraction of robust and repeatable salient points

suitable for medical image retrieval. The proposed retrieval method achieved an average retrieval rate of 65% for 20 retrieved images. This result compares well among with those in related work in the literature.

In [144], Deselaers et al. evaluated the performance of various features for medical image retrieval. The results were reported by using the complete IRMA database. An average error rate of 27.7 % was reported by using SIFT-based local features. Gabor-based features achieved an average error rate of 37.9 %. In [145], Liu et al. used DCT-based salient points and Gabor features for content-based medical image retrieval. The authors used a database of 1000 medical images. The accuracy of the system was evaluated by using 50 query test images. An average retrieval rate of 54 % was achieved for 40 retrieved images.

In [155], Rao et al. proposed a content-based image retrieval system using the Contourlet transform. Global features were used for the retrieval task. Experiments were conducted by using a database of 750 face images. The authors reported an average retrieval rate of 68.95 % for 15 retrieved images. Compared to the work of Rao et al. in [155], our method seems to be more suitable for medical applications, for the following reasons:

- The proposed method uses the Non Subsampled Contourlet Transform (NSCT) for low-level image analysis. The NSCT is shift invariant, so it allows for the extraction of rich and discriminating information from different sub-bands.
- The proposed method relies on salient points and local descriptors. In medical retrieval applications, localized features are more suitable than global features because of the variations in image quality and the appearance of anatomical structures.

6.7 Conclusions

In this chapter, we presented new methods for two important medical imaging problems: the segmentation and retrieval of anatomical structures in digital radiographs. For image segmentation, we were particularly interested in the segmentation of vertebral structures in x-ray images. Our main focus was to develop an efficient segmentation method that can deal with both normal and pathological vertebrae. For image retrieval applications, we were interested in the retrieval of radiograph images with various normal and pathological anatomical structures.

The following summarizes the contributions presented in this chapter:

- First, at the core of the proposed segmentation and retrieval algorithms is a novel appearance-based salient point detection and matching method. The new method is based upon the Non Subsampled Contourlet Transform (NSCT), a truly 2D multi-scale and multi-directional image analysis tool. Contrary to the majority of salient point detectors in the literature, our detector integrates prior local appearance knowledge in the detection process. In comparison to related methods in the literature, our method has proven successful in detecting stable and relevant salient points in medical images. Rich and discriminating information was extracted by using local descriptors from the non-subsampled contourlet sub-bands.
- Second, we presented a novel method for the segmentation of lumbar vertebrae by using shape-driven salient point matching in the Contourlet domain. The proposed method integrates low-level and high-level approaches. At the lower-level, the contourlet-based salient point detector is used to detect evidence of the object of interest in the test image. At the higher-levels, Contourlet-based local appearance profiles and multi-scale shape prior are used to drive the segmentation process. The performance of the proposed method was thoroughly evaluated in section 6.6.
- Third, we presented a novel content-based medical image retrieval system using Contourlet-based salient point detectors. The proposed method retrieves images of similar classes by using salient points and Contourlet-based local descriptors. The proposed method was successful when applied to the IRMA database.

7. Simulation of Pathological Deformations Using Localized Modeling of Shape and Texture Variability

7.1 Introduction

This chapter presents a novel method for the sparse modeling of texture variability in medical images. The chapter also describes a general framework for the simulation of pathological deformations in x-ray images. Recently, there has been a growing interest in the concept of sparsity and its applications in the field of computer vision. In addition to the direct application of sparsity in signal compression, the role of sparsity in machine learning has lately emerged as a promising approach. Traditionally, texture appearance models are constructed by using a global PCA-based basis. A classical example is the Active Appearance Model (AAM) proposed by Cootes [37]. However, it is now accepted that most pathologies are often related to local variations in both texture and shape.

The first part of this chapter introduces a novel sparse texture appearance model based upon Contourlet transform and Independent Component Analysis (CTICA-AM). A sparse texture representation is obtained by using the Contourlet transform. Contourlets are suitable for the representation of higher-order singularities in medical images because of the multi-scale, directionality, and anisotropy of Contourlets' basis. Localized texture variations are captured by using ICA-based modeling in the compressed Contourlet domain. The modeling of raw pixel intensities becomes unfeasible as the resolution of the image increases. The proposed appearance model (CTICA-AM) benefits from the non-linear approximation power of Contourlets to achieve high data reduction rates, while preserving the accuracy of the statistical model. The new texture-enhanced appearance model outperforms other related wavelet-based models [62, 63]. This advantage is particularly important in modeling high-resolution images.

In the second part of this chapter, we present a general framework for the simulation of photo-realistic spine x-ray images with various pathological deformations. Within the medical imaging community, simulation of images is particularly important due to the scarcity of ground truth data, which are needed in a number of tasks such as validating of segmentation and registration algorithms. Statistical models are always challenged by the

curse of dimensionality where the size of the available training set is far smaller than the dimension of the problem. The proposed synthesis framework integrates the new enhanced texture appearance model and the multi-scale shape model presented in chapter four.

The rest of the chapter is organized as follows: Section 7.2 presents the details of the novel Contourlet enhanced texture appearance model. Section 7.3 describes a general image synthesis framework for spine x-ray images. Our experiments and results are presented in section 7.4. Finally, conclusions are drawn in section 7.5.

7.2 Sparse texture appearance model using Contourlet-ICA

This section presents a novel sparse texture appearance model based upon the Contourlet transform and Independent Component Analysis. The proposed model is capable of extracting spatially localized texture variations favouring sparsity. The model exploits the following concepts: (i) the ability of Contourlets to represent higher-order singularities, (ii) the sparsifying power of Contourlets, and (iii) the locality of the ICA basis. Figure 7.1 summarizes the building blocks of the new Contourlet enhanced texture appearance model (CTICA-AM).

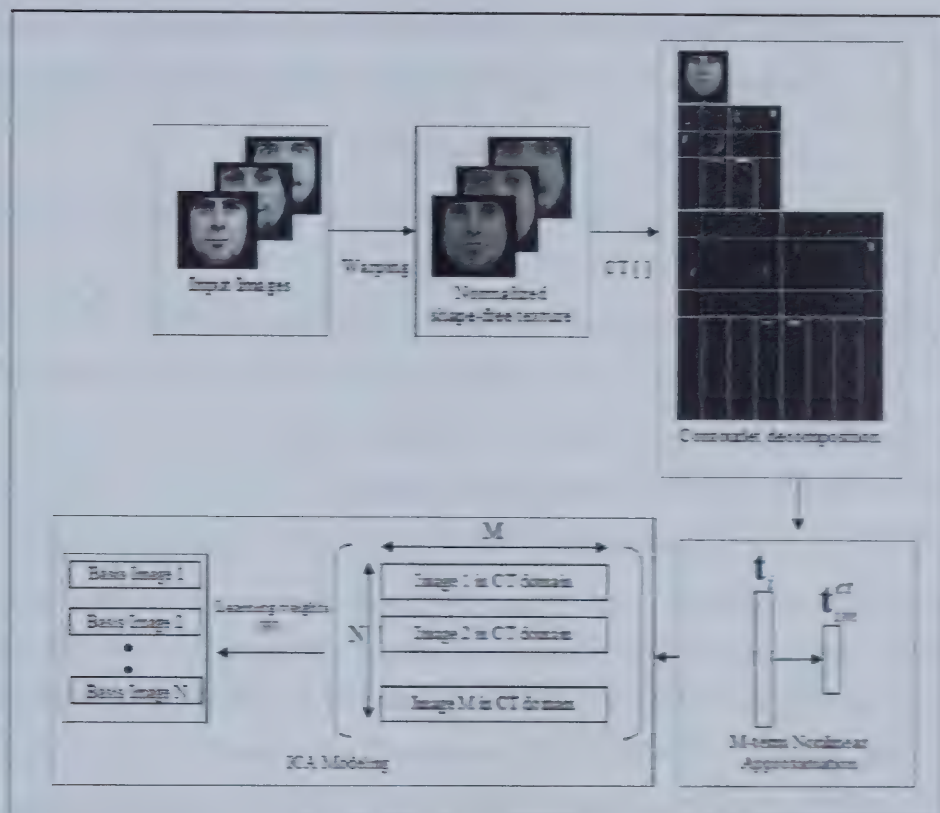


Figure 7.1 Block diagram for the Contourlet enhanced texture appearance model (CTICA-AM)

7.2.1 Contourlet decomposition

Transforms try to represent a signal as sparse as possible (i.e., with few coefficients). Traditionally, sparsity has been exploited for compression applications. However, the role of sparsity in machine learning has received much interest over the past few years. The rationale is that statistical modeling in a sparse domain can reduce the dimensionality of the problem and achieve accurate results by using limited training set. A number of transforms can be used to produce a sparse representation of an image. Among various transforms, wavelets and wavelets packets have received much interest over the past decades. However, it is now accepted that wavelets are optimal only in representing 1D signals. Wavelets fail to efficiently represent higher-order singularities due to wavelets' inability to "see" the smoothness along contours. Contourlets were introduced by Do, et al. as a geometrical second generation of wavelets that can better deal with higher-order singularities in images [99]. This advantage is attributed to the appealing properties of the Contourlet basis such as multi-scale, multi-directionality, and anisotropy.

We start with a training set of images, each with K corresponding landmark points marking the structure of interest. Let the n^{th} shape vector be represented as:

$$S_n = [x_{n1}, \dots, x_{nK}, y_{n1}, \dots, y_{nK}]^T, \quad (7.1)$$

where x and y are the x - and y -coordinates, respectively.

Prior to texture modeling, we use the Generalized Procrustes method to normalize the shapes with respect to translation, rotation, and scaling. A reference shape (S_o) is taken to be the Procrustes mean shape. Next, a set of shape-free texture images is created by warping the convexhull of the landmarks in each training sample to the reference shape (S_o). This warping is done by using the piece-wise affine warping method [39]. This step establishes the correspondence among the images in the training set and allows for modelling texture variability. We use two sets of images to evaluate the proposed texture appearance model. The first set of images is a subset of the IMM face database created by the department of Informatics and Mathematical Modeling at the Technical University of Denmark [156]. The second set of images is a set of spine x-ray images obtained from the National Library of Medicine database [108]. Figure 7.2 shows examples of the face and spine images used in our experiments.

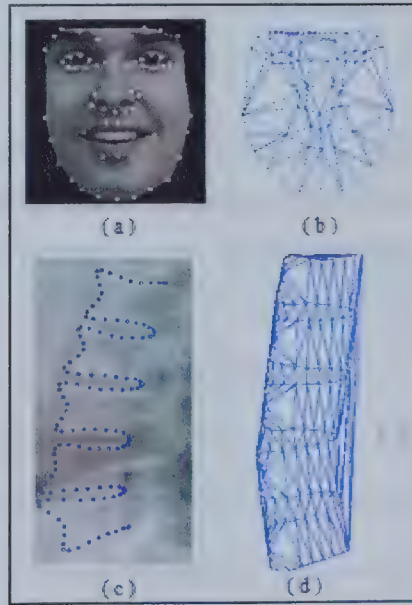


Figure 7.2 (a) Face image with landmarks, (b) triangulated reference face shape, (c) Spine image with landmarks, (d) triangulated reference spine shape.

Given a training set of shape-free texture images, we perform a multi-scale multi-directional image decomposition by using the Contourlet transform. We use a 4-level Contourlet transform with 0, 4, 8, and 16 directions, respectively.

Let $\{\mathbf{g}_i\}_{i=1..M}$ be the set of vectors of pixel intensities sampled at L locations for the texture images in the training set:

$$\mathbf{g}_i = [\mathbf{g}_{i1}, \mathbf{g}_{i2}, \mathbf{g}_{i3}, \dots, \mathbf{g}_{iL}]^T. \quad (7.2)$$

$\{\mathbf{t}_i\}_{i=1..M}$ is the set of vectors of concatenated Contourlet coefficients resulting from the decomposition of the shape-free texture images in the training set:

$$\mathbf{t}_i = \mathbf{C}_T[\mathbf{g}_i], \quad (7.3)$$

where $\mathbf{C}_T[.]$ is the contourlet transform operator. M is the size of the training set.

7.2.2 Non-linear approximation

We exploit the approximation power of Contourlets to tackle the problem of modeling high-resolution images. We use this power to achieve high reduction rates for the raw texture images prior to statistical modeling. Statistical modelling is then performed on the compressed Contourlet texture space.

The non-linear approximation power of a transform is evaluated in terms of its partial reconstruction error. The m -term approximation of an image is obtained by reconstructing the transformed image by using the highest m coefficients in the transformed domain. According to [39], Contourlets exhibits optimal behaviour in approximating smooth objects with discontinuities along the C^2 curve. The mean square reconstruction error (MSE) for m -term approximation grows with m^{-1} by using the wavelet basis. For the Contourlet basis, the reconstruction error grows with $\log(m^3)m^{-2}$:

$$\|\mathbf{t} - \mathbf{t}_m^{wt}\|_2 \propto m^{-1}, \quad m \rightarrow \infty \quad (7.4)$$

$$\|\mathbf{t} - \mathbf{t}_m^{cf}\|_2 \propto \log(m^3)m^{-2}, \quad m \rightarrow \infty, \quad (7.5)$$

where $\mathbf{t}_m^{wt}$, and $\mathbf{t}_m^{ct}$ are the m -term approximation using the wavelet and contourlet basis, respectively.

For a training set of $\mathbf{P}$ images, we have a set of vectors of concatenated Contourlet coefficients $\{\mathbf{t}_i\}_{i=1..P}$. We select a subset of the Contourlet coefficients by choosing the m coefficients that best approximate the texture vectors, while preserving the per-coefficient variance over the training set. Let $\mathbf{C}$ denote a modified identity matrix reflecting the indices of the coefficients to be retained. The compressed Contourlet-based representation of images in the training set is then given by $\{(\mathbf{t}_m^{ct})_i\}, i = 1..P$, where

$$(\mathbf{t}_m^{ct})_i = \mathbf{C}\mathbf{t}_i. \quad (7.6)$$

7.2.3 ICA modelling in compressed Contourlet domain

Next, we use Independent Component Analysis (ICA) to model the texture variability in the compressed Contourlet domain. Pathologies are typically localized in both shape and texture. We propose to use ICA to extract spatially localized texture variations favouring sparsity.

Independent Component Analysis (ICA) is viewed as a generalization of Principal Component Analysis (PCA). On the one hand, PCA can learn only second-order dependencies among the image pixels. On the other hand, ICA can also learn higher-order dependencies. In this work, we use ICA to extract localized modes of texture variations in the compressed Contourlet domain. We use the FastICA algorithm to find the independent sources of variations in the training set. The FastICA algorithm is an efficient method that finds independent components by maximizing non-Gaussianity in the data matrix [117].

The ICA model is defined as follows:

$$\begin{aligned} X &= AS \\ Y &= WX \quad , \end{aligned} \quad (7.8)$$

where X is a matrix of observed mixed signals, and S is a matrix of statistically independent source signals. A and W are the mixing and de-mixing matrices, respectively. Y is a matrix with the estimated source signals.

The rows of the X matrix contain the vector representation of the texture images in the training set. The rows of the Y matrix contain the estimated statistically independent basis images. A visualization of the ICA texture appearance model is given in Figure 7.3.

Rather than performing the ICA modeling in the pixel-intensity space, we perform the ICA modeling in the compressed Contourlet domain. Applying ICA to the selected contourlet coefficients has the following advantages. First, doing so improves the performance of the ICA algorithm since projecting the data matrix onto a sparse domain increases the independence of the sources. Second, doing so increases the robustness to noise and missing information by removing the noisy components through the Contourlet approximation step.

Let $C_T[\cdot]$ be the linear contourlet transform; then the ICA model can be directly re-written in the contourlet domain as

$$C_T[X] = AC_T[S] \quad (7.9)$$

$$C_T[Y] = WC_T[X] \quad (7.10)$$

The Contourlet-ICA basis matrix is then defined as $\Phi_{CT-ICA} = C_T[Y]$. $C_T[Y]$ is a matrix whose rows contain the statistically independent basis in the contourlet domain.

The proposed model can be used to generate new texture images. We use the Contourlet-ICA model to predict selected Contourlet coefficients. We insert average coefficient values at the positions of all truncated coefficients.

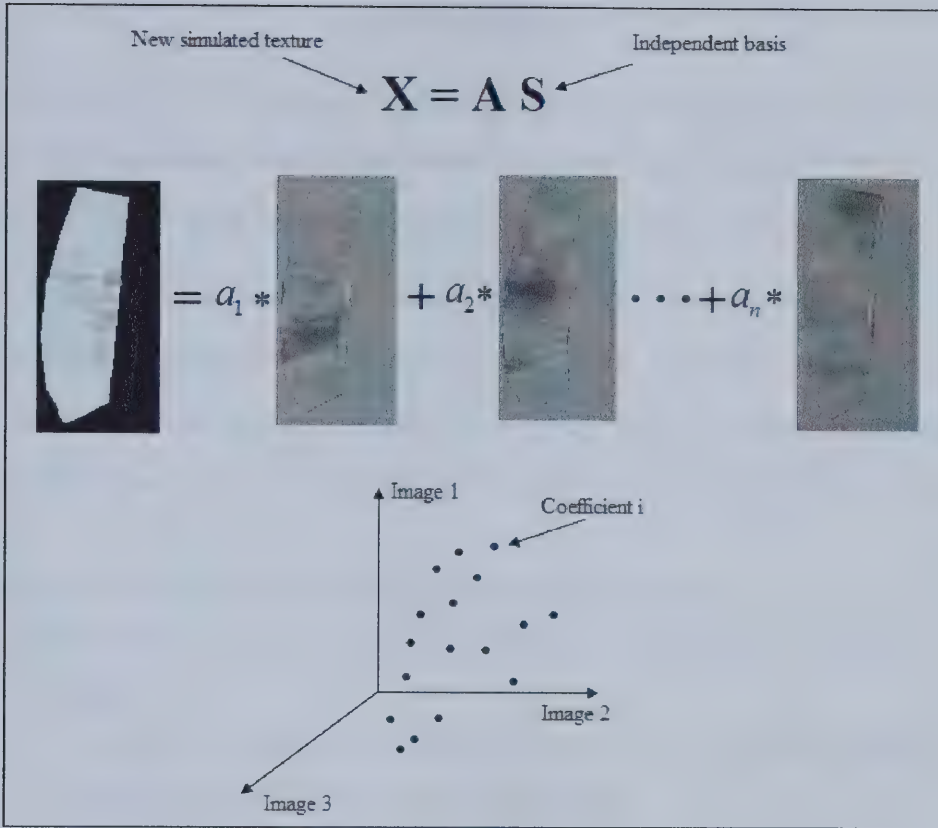


Figure 7.3 Enhanced texture appearance modeling using ICA basis in Contourlet domain

7.3 Image synthesis framework

This section presents the details for a novel image synthesis framework to generate simulated spine x-ray images with various pathological deformations. The synthesis of medical images is important for validating segmentation and registration algorithms. Our goal is to create a simulation framework capable of imposing certain pathology-related deformations such as disc narrowing, and vertebral fractures. We focus mainly on the synthesis of spine x-ray images. However, the proposed framework can also be used to synthesize images of other anatomical structures.

The new framework integrates the shape and texture generative models proposed earlier in this thesis, along with an example-based pose model.

The following presents the details of the proposed framework:

A grey level image I is viewed as a mapping from R^2 space to the range of possible gray levels:

$$I: R^2 \rightarrow [0, 255]. \quad (7.11)$$

Let $\{I_i, S_i\}, i = 1..N$ be a set of N training images. I_i denotes the input image. S_i is a vector of the landmarks outlining the contour of the anatomical structure of interest.

Let $\{I_o, S_o\}$ denote the reference image and the corresponding reference contour.

Let A_i denote the affine transformation aligning the reference shape (S_o) to the shape S_i . This transformation is found by using Procrustes Alignment [38].

Let $W(.)$ denote the warping operator. The warping function is estimated by using the piece-wise affine warping method [39], which is efficient and allows for inverse warping.

An image I is characterized by the following three components:

- The texture component, which determines the appearance of the structure of interest.
- The shape component, which outlines the contour of the structure of interest.
- The pose of the structure of interest in the image.

Using the training set of images $\{I_i, S_i\}, i = 1..N$, we learn the variability in the pose, shape, and texture spaces. In each of these spaces, the image is viewed as a point whose coordinates constitute a parameter vector. Let $\mathbf{a}_P$, $\mathbf{a}_S$, and $\mathbf{a}_T$ represent the parameter vectors of the pose, shape, and texture, respectively.

We learn the shape space by using the multi-scale shape model presented in chapter four. The generative Contourlet enhanced appearance model (CTICA-AM) is used to model the texture space. An example-based pose model is used to represent variations in translation and scaling within the input data set.

- Multi-scale shape model

We model variations in the shape of the anatomical structure by using the multi-shape model presented in chapter four. Using wavelet-packets and Independent Component Analysis, we extract multi-scale modes of deformations. Due to the locality of the Wavelet-ICA basis functions, the model can be used to selectively impose various

pathology-related deformations such as disc-narrowing, osteophytes, and vertebral fractures.

We start with a training set of landmarks outlining the structure of interest in the images. Let Φ_{WICA} be the matrix of Wavelet-ICA basis set described in chapter four. The multi-scale generative shape model can then be written as follows:

$$S = S_o + \alpha_s \Phi_{\text{WICA}} . \quad (7.12)$$

A new shape (S) is represented as a reference shape (S_o) and a weighted sum of the wavelet-ICA multi-scale basis functions. The weighting vector (α_s) is used to impose a certain localized deformation. The multi-scale shape model also includes an inverse transformation step from the wavelet-packet domain to the spatial domain. This step is eliminated from equation 7.12 for clarity.

- Texture appearance model

We use the Contourlet enhanced texture model presented in section 7.2 to model localized texture variations. We start with a training set of spine x-ray images. Each image has a number of landmarks outlining the contour of the lumbar spine. We also add a number of border landmarks such that the whole image is encoded into the texture model. A set of normalized shape-free texture images, $\{T_{im}^i\}, i=1...N$, is obtained by warping the images to the reference shape using the piece-wise affine warp. This procedure brings the pixels into correspondence. Next, we use the enhanced appearance model described in section 7.2 to extract a set of Contourlet-ICA basis functions (Φ_{CTICA}).

A new texture vector (T) is generated by the summation of the reference texture vector (T_o) and the weighted sum of the Contourlet-ICA basis functions (Φ_{CTICA}):

$$T = T_o + \alpha_T \Phi_{\text{CTICA}} , \quad (7.13)$$

where α_T is the weighting vector used to generate a new texture image.

- Example-based pose model

We generate new poses of the spine by using an example-based pose model. Given a training set of lumbar spine images with various poses ($\{I_i, S_i\}, i = 1..N$), we select a number of proto-type poses, which are used to represent the pose space. We interpolate these proto-type pose examples to generate a new pose of the spine. Let $\{\tilde{S}_o^i, i = 1..M\}$ be the set of proto-type poses of the reference shape obtained by affine warping the reference shape to the selected pose examples. $\tilde{S}_o^i$ and S_i have the same pose:

$$\tilde{S}_o^i = A_i S_o. \quad (7.14)$$

We use interpolation techniques to synthesize a new pose within the M-1 dimensional pose space. A new pose for the reference shape ($S_{o,p}$) is defined as a weighted sum of the selected proto-type poses:

$$S_{o,p} = \sum_i \alpha_p \tilde{S}_o^i. \quad (7.15)$$

Figure 7.4 shows an example of a two-dimensional pose space using four proto-type poses $\{\tilde{S}_o^i, i = 1, 2, 3, 4\}$. We use bilinear interpolation to synthesize a new pose for the reference shape ($S_{o,p}$) at the (x, y) position of the pose space:

$$S_{o,p} = \sum_{i=1}^4 \alpha_p(i) \tilde{S}_o^i, \quad (7.16)$$

where:

$$\alpha_p(1) = (1-x)(1-y), \alpha_p(2) = (x)(1-y),$$

$$\alpha_p(3) = (1-x)(y), \alpha_p(4) = (xy), \text{ and } x, y \in [0, 1].$$

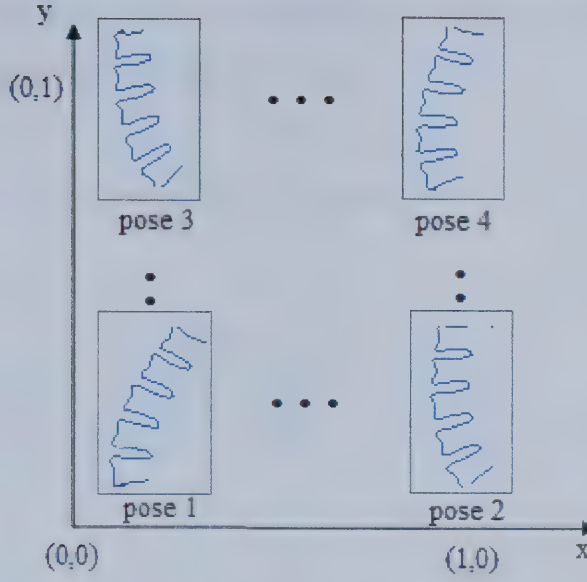


Figure 7.4 Example of two-dimensional pose space using four proto-type poses

Figure 7.5 shows the building blocks for the proposed image synthesis framework. An x-ray image with a certain imposed pathology can be synthesized by using the following procedure:

- i. Starting with the reference shape (S_o), we use the basis of the multi-scale shape model to impose the desired pathological deformation. The outcome is a new shape (S_n) (see Equation 7.12).
- ii. A new texture appearance (T_n) is then generated by using the basis set of the Contourlet enhanced appearance model (Φ_{CTICA}) and the texture parameter vector ($\mathbf{a}_T$) (see Equation 7.13).
- iii. We use the pose parameter vector ($\mathbf{a}_p$) to generate a new pose for the reference shape ($S_{o,p}$). We then find the transformation matrix (A_p) that maps S_o to $S_{o,p}$.
- iv. The new shape (S_n) is then warped to the new position by using the transformation matrix (A_p):

$$S_{n,p} = A_p S_n. \quad (7.16)$$

$S_{n,p}$ represents the generated shape with the new pose.

- v. A new image is finally synthesized by rendering the synthesized shape-free texture to the new shape vector $S_{n,p}$:

$$I_n = W(T_n, S_{n,p}), \quad (7.17)$$

where $W(.)$ is the operator for the warping function.

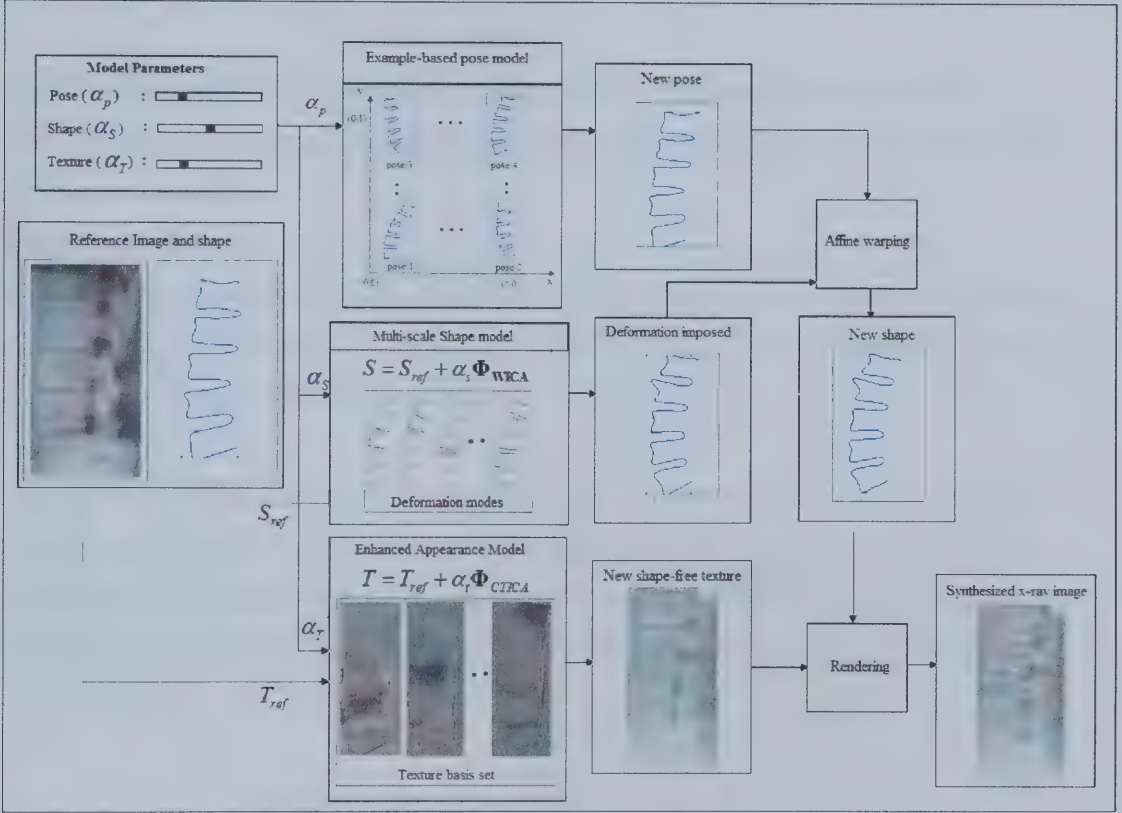


Figure 7.5 Block diagram of proposed spine x-ray image synthesis framework

7.4 Experiments and Results

This section presents the details for the experiments conducted to evaluate the methods presented in this chapter. We use two datasets of images. The first one is a dataset of 30 x-ray images of the lumbar spine, with their corresponding ground truth segmentations [108]. The second one is a dataset of 40 images of faces for different individuals with various expressions [156].

7.4.1 Evaluation of the Contourlet enhanced texture appearance model

In this section, we evaluate the performance of the proposed Contourlet enhanced texture appearance model.

Approximation power

We assess the approximation power of the proposed model for the x-ray images and the facial images. Our results are compared to those from a wavelet-based texture appearance model proposed by Stegman et al. in [62, 63].

For a desired reduction ratio (RR), we calculate the corresponding number of coefficients to keep (M). Next, we obtain an M-term approximation of the image in both the contourlet and wavelet domains. We compare the quality of the reconstructed signal by using the Signal-to-Noise Ratio (SNR). This step is repeated for various reduction ratios:

$$SNR = 10 \log_{10} \frac{\sum_x \sum_y I(x, y)^2}{\sum_x \sum_y (I(x, y) - I_M(x, y))^2}, \quad (7.18)$$

Where $I(x, y)$ the original is image and $I_M(x, y)$ is the image reconstructed using M coefficients in the Contourlet domain.

Table 7.1 shows the SNR for the contourlet-based representation of the spine x-ray images versus the wavelet-based representation. In general, Contourlets tend to achieve a higher SNR, especially at higher reduction ratios. At a reduction ratio of 1:40, the Contourlet-based representation shows an increase of 11.9 % in the SNR. As the reduction ratio increases to 1:200, the contourlet-based representation shows an increase in the SNR of 86.3 %. This finding agrees with the results obtained from the visual inspection of the reconstructed images in Figure 7.6. The wavelet-based representation

deteriorates rapidly at higher reduction rates. At a reduction ratio of 1:200, the Contourlet-based representation appears to better capture the details of the image.

Table 7.1 Reconstruction of spine x-ray images at different reduction ratios for the Contourlet-based and Wavelet-based appearance models.

Reduction ratio	1:200	1:125	1:100	1:40
SNR for Contourlet-based model	10.66 db	12.91 db	13.04 db	14.85 db
SNR for Wavelet-based model	5.72 db	7.41 db	10.16 db	13.26 db

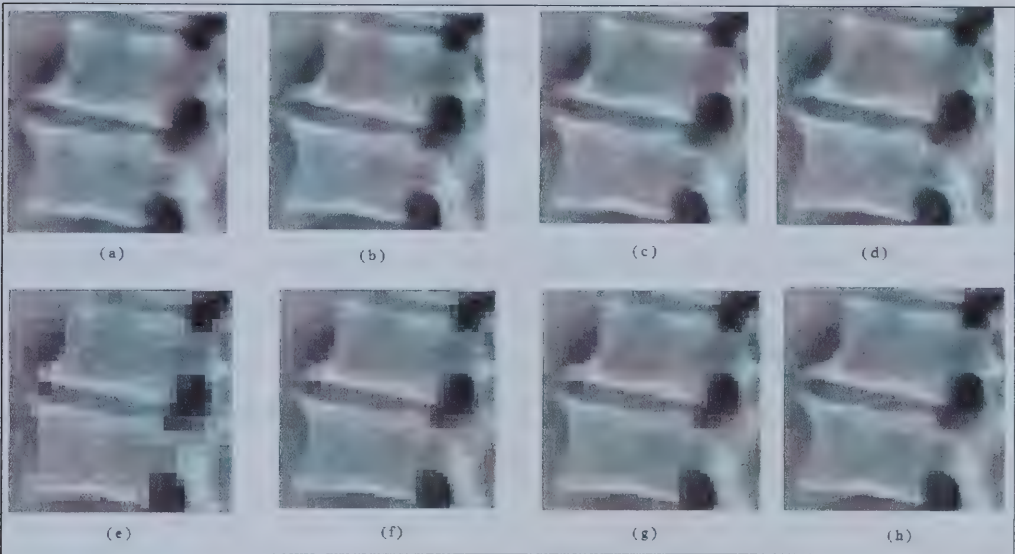


Figure 7.6 Non-linear approximation of a spine x-ray image using Contourlets (top row), and Wavelets (bottom row). Reduction ratios: (a , e) 1: 200 (b , f) 1:125 (c , g) 1:100, (d , h) 1:40.

Similar results were obtained when our proposed model was applied to the face images. results are shown in Table 7.2 and Figure 7.7.

Table 7.2 Reconstruction of facial images at different reduction ratios for the Contourlet-based and Wavelet-based appearance models.

Reduction ratio	1:200	1:125	1:100	1:40
SNR for Contourlet-based model	6.87 db	8.0 db	8.86 db	11.73 db
SNR for Wavelet-based model	4.97 db	5.31 db	5.55 db	9.96 db



Figure 7.7 Non linear approximation of facial image by using Contourlets (Top row), and (Wavelets).
Reduction ratios: (a , e) 1: 200 (b , f) 1:125 (c , g) 1:100, (d , h) 1:40

Locality of Contourlet-ICA basis

Figure 7.8 compares the basis sets of the proposed appearance model (CTICA-AM) and the PCA-based basis set used in the Active Appearance Model (AAM). The proposed model produces more localized basis functions than the global PCA basis. This property is further exploited to impose localized texture variations. Figure 7.9 shows examples of facial images generated by using the new Contourlet-ICA texture appearance model. Thanks to the locality of the basis functions, we are able to selectively impose various localized deformations.



(a)



(b)

Figure 7.8 Examples of basis sets for different texture appearance models. (a) PCA-based Active Appearance Model, (b) Contourlet enhanced appearance model.



Figure 7.9 Examples of localized texture variations obtained by using the new sparse texture model.

Texture Reconstruction

We study the ability of the proposed appearance model to reconstruct texture appearance examples by using a limited-sized training set. We conduct the leave-one-out experiment where by the appearance model is trained by using N-1 training images, and the left-over image is used for testing. This procedure is repeated for all the N images in the training set. The average reconstruction error (also referred to as the generalization error) is then calculated as the mean intensity error per pixel (for a 256 gray level image). This procedure is repeated for N= [10 20 30 40]. Figure 7.10 compares the performance of the proposed appearance model (CTICA-AM) to that of the Active Appearance Model (AAM). The new texture appearance model is shown to have a lower reconstruction error for training sets of different sizes. On average, the proposed Contourlet-ICA appearance model achieves a decrease of 17.62 % in the texture reconstruction error compared to the AAM model.

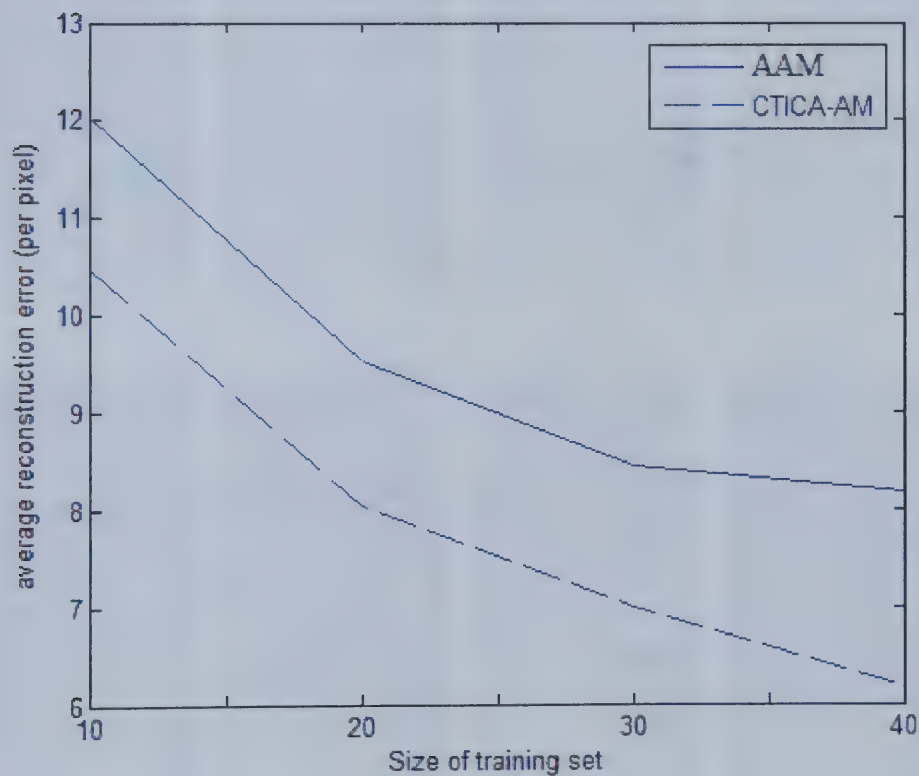


Figure 7.10 Reconstruction error of the new texture appearance model for different sizes of the training sets.

7.4.2 Evaluation of the image synthesis framework

This section evaluates the performance of the complete image synthesis framework. We use a training set of spine x-ray images and the corresponding ground truth segmentation. Figure 7.11 shows examples of synthesized x-ray images where various localized pathology-related shape deformations were imposed. Figure 7.11(h) shows example of disc space narrowing between the L1 and L2 vertebrae. Figure 7.11(b,c) show signs of osteophyte pathology developed for the L2 vertebra. Figure 7.11(f) shows a fracture that was imposed on the L4 vertebra. Figure 7.12 shows examples of synthesized x-ray images with localized variations in texture appearance. The shape and pose parameters are kept constant.

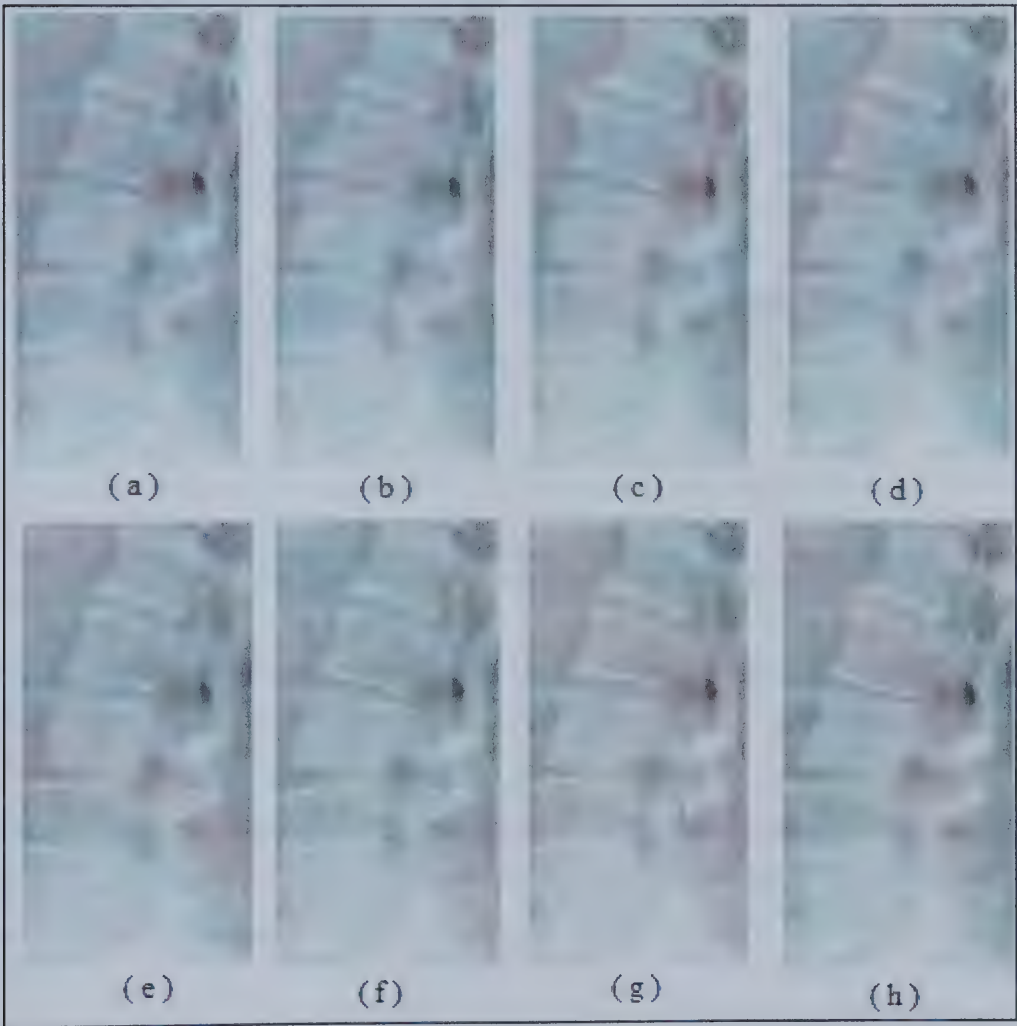


Figure 7.11 Examples of simulated x-ray spine images with various localized shape deformations being imposed (The texture and pose parameters are kept constant).

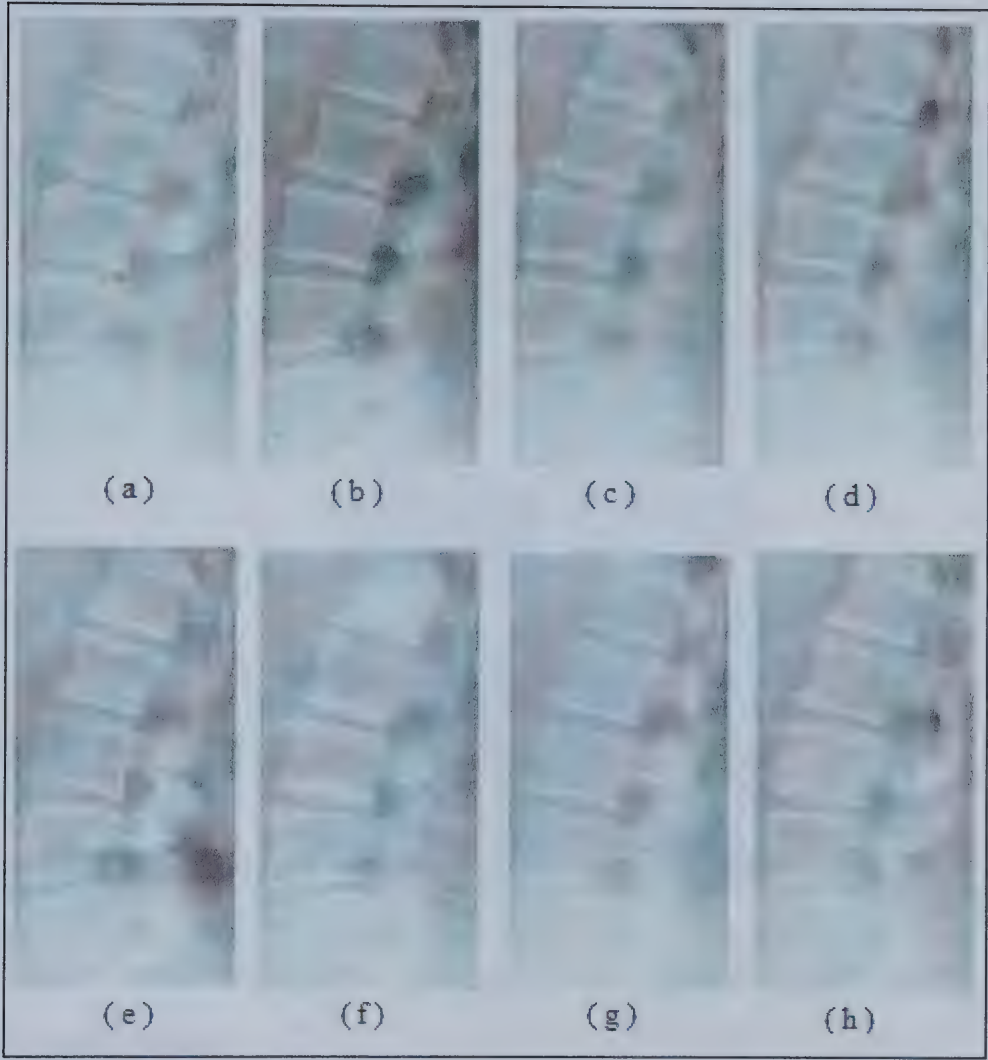


Figure 7.12 Examples of simulated x-ray spine images with various texture variations being imposed (The pose and shape parameters are kept constant).

By varying the model parameters ($\alpha_p, \alpha_s, \alpha_T$), we can simulate photo-realistic x-ray images with different degrees of pose, shape and texture variations. The outcome is a set of synthesized images and the corresponding ground truth segmentations. Figure 7.13 shows examples of simulated x-ray images and their ground-truth segmentations. By using the proposed framework, one can create an enlarged training set with specific shape and texture deformations.

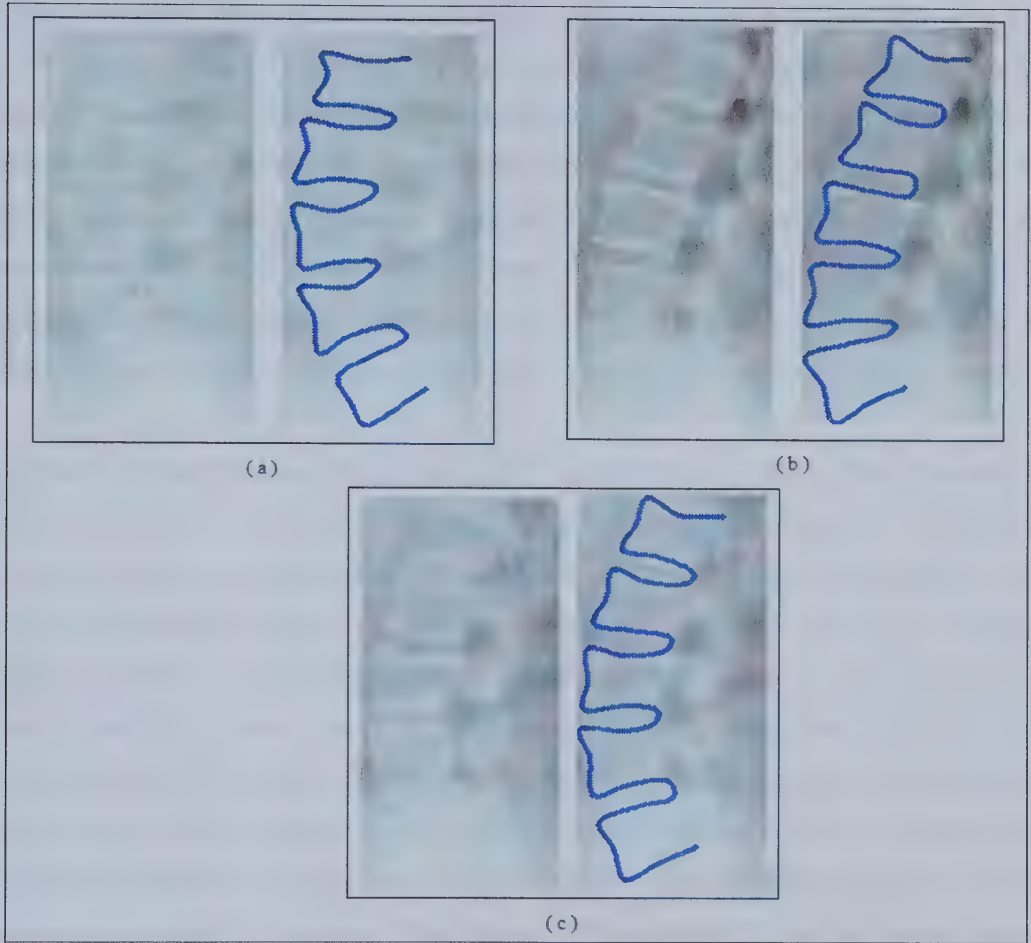


Figure 7.13 Examples of simulated spine x-ray images and the corresponding segmentations.

7.5 Conclusions

This chapter presented new methods for the simulation of pathological deformations in x-ray images by using localized models of shape and texture.

First, we presented a novel sparse texture appearance model based upon the Contourlet transform and Independent Component Analysis (ICA). The proposed appearance model benefits from the non-linear approximation power of Contourlets to achieve high data reduction rates, while preserving the accuracy of the statistical model. Our experimental results demonstrated significant performance gains for the proposed model in comparison to related work in the literature, particularly for high reduction rates. Acceptable results are obtained with reduction rates as high as 1:200. Such results are important in modeling high-resolution images where modeling of raw pixel intensities becomes unfeasible due to storage and computational requirements. The model is also capable of capturing localized texture variations through ICA modeling in the compressed Contourlet domain.

Second, we attempted to tackle the challenges related to the scarcity of ground truth data in medical imaging. We presented a general framework for the synthesis of photo-realistic spine images with various pathological deformations. The proposed framework integrates the shape and texture models developed in this thesis. We demonstrated the ability of the framework to generate x-ray images with various pathology-related shape and texture variations. Throughout this chapter, we focused mainly on the simulation of images of the human lumbar spine. However, the same framework can be used to simulate images of other anatomical structures.

The new synthesis framework compares well with related methods in the literature. In [66, 58], the authors used image morphing for image synthesis applications. New images are generated as linear combinations of proto-type examples. On the other hand, the proposed framework generates a new image by using multi-scale statistical models of shape and texture. This framework has the following advantages: (i) better removal of linear dependencies among the proto-type examples, and (ii) better extraction of complex variation modes. In [58, 64, 67, and 68], various PCA-based image synthesis methods were presented. In comparison to the PCA-based image synthesis approaches, the proposed framework can capture a wider range of localized deformations for the anatomical structure of interest. Using the new framework, we can synthesize images with specific pathology-related shape and texture variations. The proposed framework can be used to validate segmentation and registration algorithms.

8 Summary and Future Work

8.1 Summary

This thesis dealt with a number of challenges related to medical imaging and model-based computer vision. We presented novel statistical models of shape and texture to improve the support of various medical practices. Within a biologically-inspired framework, we employed these models in two spine-related tasks: the segmentation and quantification of vertebral structures in medical images.

The following summarizes the contributions of this thesis.

Multi-scale shape modelling

In chapter four, we attempted to overcome some of the challenges associated with statistical modelling in high-dimensional space. We presented a novel multi-scale statistical shape model using the concepts of sparsity and statistical independence to extract multi-scale localized modes of deformations. Within a best-basis selection framework, the proposed method benefits from the multi-scale and sparsifying nature of wavelet packets and the ability of Independent Component Analysis (ICA) to capture higher-order statistics. These benefits allow for the construction of a localized non-linear shape model with clinically relevant modes of deformation. Experiments were conducted to evaluate the performance of the new model in comparison to the Point Distribution Model (PDM) and other related methods in the literature.

Localized vertebral shape analysis and classification

The thesis also presented a novel framework for the analysis and classification of vertebral structures by using pathology-related Wavelet-ICA shape parameters. Within a statistical framework, we were able to establish a statistically relevant relationship between the Wavelet-ICA modes of deformations and clinical information. We also presented a novel method for multi-vertebrae classification using Wavelet-ICA parameters and Fisher classifier. Thanks to the locality of the features, the new classification scheme is able to assess the condition of several vertebrae simultaneously. Experiments were conducted to evaluate the performance of the proposed vertebral quantification methods. Our results show that the proposed framework achieves a higher performance than the related frameworks in the literature.

Segmentation and retrieval of anatomical structures in x-ray images

We also presented a novel method for the segmentation of vertebral structures in x-ray images. The new method incorporates the following: (i) a novel salient point detection method in the Non Sub-sampled Contourlet domain, (ii) Contourlet-based local appearance profiles for matching salient points, and (iii) a novel multi-scale shape prior used to drive the segmentation process. Experiments were conducted to evaluate the accuracy of the new segmentation method.

In addition to segmentation, we also used the proposed contourlet-based salient point detection method in a content-based image retrieval task. The proposed method benefits from the robustness and stability of the extracted salient points, and the rich Contourlet descriptors. We obtained promising results when we applied our method to the IRMA (Image Retrieval in Medical Applications) database [125]. This work represents the first contribution from the Medical group at the University of Alberta to the IRMA project.

Enhanced modelling of texture appearance

Finally, we tackled the challenges related to the accurate modeling of texture and scarcity of training data in medical imaging. We presented a new sparse texture appearance model based upon the Contourlet Transform and Independent Component Analysis (ICA). The new model benefits from the non-linear approximation power of Contourlets to achieve high reduction rates. The statistical independence of Independent Component Analysis (ICA) was employed to capture localized texture variations. We presented a general framework for the synthesis of photo-realistic spine images with various pathological deformations. The proposed framework integrates the shape and texture models developed in this thesis. Experiments were conducted to evaluate the performance of the Contourlet enhanced texture model and the image synthesis framework.

8.2 Future research directions

The following discusses some of the remaining challenges and future research directions related to the work presented in this thesis.

Role of sparsity in medical imaging and computer vision

Throughout this thesis, we tried to bring recent advances in computational harmonic analysis to the field of medical imaging. We were motivated by two main factors: (i) our understanding of the Human Visual System and the role of sparsity and multi-scale representation in early vision stages, and (ii) recent advances in the field of harmonic analysis and the prospects they bring to the fields of computer vision and medical imaging. The sparsity of a representation has proven useful in classical signal processing problems such as data compression. In medical imaging and computer vision, we are typically concerned with more sophisticated tasks related to image interpretation and inference. Hence, there is a growing demand for a rigorous investigation of the prospects of sparsity in various medical imaging and computer vision tasks.

Statistical modeling of shape and texture

- In an attempt to overcome the limitations of PCA-based models, this thesis focused on using Independent Component Analysis (ICA) for the statistical modeling of shape and texture variations of anatomical structures. In the future, one might investigate the performance of other non-linear generative models.
- In chapter seven, we described a novel Contourlet enhanced texture appearance model. An interesting direction for future investigation would be to use this model for the segmentation of structures in high-resolution images.
- The image synthesis framework presented in chapter seven is capable of generating images with specific pathological deformations. A potential extension to this work is to use the proposed image synthesis framework to validate different segmentation and registration algorithms.

Quantification of vertebral structures

- In chapter five, we presented a framework for vertebral analysis and classification using parameters of the multi-scale shape model. A future

extension to this framework would be to include the parameters of the localized texture appearance model presented in chapter seven. Pathologies tend to be localized in both shape and appearance. Thus, we anticipate that the integration of localized shape and texture parameters would enhance the overall performance of the classification scheme.

- Pathology at a specific vertebra is likely to affect other neighbouring vertebrae. Thus, a possible extension to the work presented in chapter five would be to include information about neighbouring vertebrae in the classification framework.
- Finally, the proposed quantification framework can be used to assess the condition of the human spine with respect to other pathologies such as fractures.

Content-based medical image retrieval

Over the past decade, researchers have been increasingly interested in content-based medical image retrieval. Most of the methods presented in this thesis can be very beneficial in medical image retrieval tasks.

- In chapter four, we described a new multi-scale shape model with clinically-relevant modes of deformations. A promising application for this model is to use the localized model parameters to develop an efficient pathology-related image retrieval system.
- In chapter six, we proposed a novel Contourlet-based salient point detection method for the retrieval of images with different classes. Experiments were conducted by using a subset of the IRMA database. In the future, we need to conduct a more thorough investigation of the prospects of this approach by using a larger database. Also, one might consider combining this approach with other image-indexing techniques in the literature.

A Unified registration, segmentation, and modeling approach

The problem of non-rigid registration is closely tied to the methods presented in this work. In this thesis, the main focus was the statistical modelling problem for segmentation and analysis applications. A possible future research direction would be to

work on a unified framework that integrates non-rigid registration methods with the modeling and segmentation methods presented in this thesis.

Extensions to other application domains

The methods developed throughout the thesis were evaluated mainly within the context of medical imaging. However, most of these methods can be applied to other domains.

- The proposed Contourlet-based salient point detection and matching algorithm will be beneficial in other computer vision applications such as object tracking.
- The multi-scale shape representation and modeling approach presented in chapter four can be used in advanced video processing techniques related to object-based predictive coding.
- The synthesis of biometric data has recently received increasing attention as this application opens the door for enhanced security. The multi-scale shape and texture models proposed in this thesis can be used for the synthesis of artificial biometric data such as the face, iris, and fingerprints.

8.3 Final Statement

This thesis was motivated mainly by the growing demand to develop computer-assisted methods for medical image interpretation and diagnosis. To achieve our goals, we presented new methods for the multi-scale statistical modeling of shape and texture. We were concerned mainly with the problems of the segmentation and quantification of vertebral structures. We hope that the contributions presented in this thesis will help medical practitioners to improve the diagnosis and tracking of diseases. Also, in a wider sense, we hope that this work will help advance the research efforts to develop efficient model-based computer vision systems.

BIBLIOGRAPHY

- [1] K. A. Walker. "Anatomy of Spinal Fractures," website for spine universe, www.spineuniverse.com, 2009.
- [2] Fact Sheet: Osteoarthritis. American College of Rheumatology, Atlanta, GA; 2006.
- [3] L. Hedlund, J. Gallagher. "Vertebral morphometry in diagnosis of spinal fracture," *Journal of Bone and Mineral Research*, vol. 5, pp.59-67, 1988.
- [4] H. Genant, C. Wu, C. Kuijk van, M. Nevitt. "Vertebral fracture assessment using a semi-quantitative technique," *Journal of Bone and Mineral Research*, vol. 9, pp. 1137-48, 1993.
- [5] Disease statistics: fast Facts. Website for the National Osteoporosis Foundation, www.nof.org/osteoporosis/stats.htm.
- [6] J. Roach. "Adolescent idiopathic scoliosis," *Orthopaedic Clinics of North America*, vol. 30, pp. 353-365, 1999.
- [7] J. Cobb. "Outline for the study scoliosis, instructional course lectures," *The American Academy of Orthopaedic Surgeons*, vol. 5, pp. 261-275, 1948.
- [8] I. Stokes, D. Aronsson. "Disc and vertebral wedging in patients with progressive scoliosis," *Journal of Spinal Disorders*, vol.14, no. 4, pp. 317-322, 2001.
- [9] N. Paragios, Y. Chen, and O. Faugeras. *Handbook of mathematical models in computer vision*. New York: Springer, 2005.
- [10] D. Kendall. "The diffusion of shape," *Advances in Applied Probability*, vol. 9, pp. 428-430, 1977.

- [11] R. Malladi, J. Sethian, B. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 17(2), pp. 158-175, 1995.
- [12] L. Staib, J. Duncan, "Boundary finding with parametrically deformable models," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 14(11), pp. 1061-1075, 1992.
- [13] F. Bookstein, *Morphometric tools for landmark data: Geometry and Biology*. Cambridge University Press, 1991.
- [14] T. Cootes, A. Hill, C. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," *Proceedings of Information Processing in Medical Imaging*, pp. 33-47, Berlin, Germany, 1993.
- [15] S. Pizer, D. Fritsch, P. Yushkevich, V. Johnson, and E. Chaney, "Segmentation, registration, and measurement of shape variation via image-object shape," *IEEE Transaction on Medical Imaging*, vol. 18(10), pp.851-865, 1999.
- [16] J. Suri, A. Farag, *Deformable Models: Biomedical and Clinical Applications*. Springer-Verlag, 1st edition, 2007.
- [17] S. Derrode, M. Charmi, F. Ghorbel, "Fourier-based Invariant Shape Prior Snakes," *Proceedings of ICASSP'06*, France, 2006.
- [18] G. Hamarneh, T. Gustavsson, "Statistically constrained snake deformations," *Proceedings of IEEE Conference on Systems, Man, and Cybernetics*, vol. 3, pp. 1610-1615, 2000.
- [19] G. Hamarneh, Tomas Gustavsson, "Constraining contour deformations using statistical apriori Knowledge of shape without requiring point-to-point correspondence," *Proceedings of Swedish Symposium on Image Analysis (SSIA-2000)*, pp. 33-36, 2000.

- [20] R. Davies, T. Cootes, and C. Taylor, "A minimum description length approach to statistical shape modelling, " In *Information Processing in Medical Imaging*, LNCS 2082, pp.50-63, Springer-Verlag, 2001.
- [21] T. Cootes and C. Taylor, "Anatomical statistical models and their role in feature extraction", *British Journal of Radiology*, vol. 77, pp. S133-S139, 2004.
- [22] P. Yushkevich, S. Pizer, S. Joshi, and J. Marron, "Intuitive localized analysis of shape variability," *Information Processing in Medical Imaging (IPMI-01)*, pp. 402-408, 2001.
- [23] G. Hamarneh, R. Abugharbieh, and T. McInerney, "Medial Profiles for Modeling Deformation and Statistical Analysis of Shape and their Use in Medical Image Segmentation," *International Journal of Shape Modeling*, vol. 10(2), pp. 187-209, 2004.
- [24] S. Osher, N. Paragios. *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Springer-Verlag, 2003.
- [25] M. Leventon, W. Grimson, O. Faugeras, "Statistical shape influence in geodesic active contours," *Proceedings of International Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 316-323, Hilton Head Island, South Carolina, 2000.
- [26] M. Stegmann, *Active Appearance Models: Theory, Extensions, & Cases*. MSc. Thesis, Technical University of Denmark, 2000.
- [27] A. Jain, Y. Zhong, M. Dubuisson-Jolly, "Deformable template models: A review," *Signal Processing*, vol. 71(2), pp.91-108, 1998.
- [28] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1(4), pp.321-331, 1987.
- [29] T. McInerney, D. Terzopoulos, *Deformable Models*. *Handbook of Medical Imaging: Processing and Analysis*, I. Bankman (Editor). Academic Press, 2000.

- [30] L. D. Cohen, "On active contour models and balloons," In Proceedings of conference CVGIP: Image Understanding, vol. 53(2), pp. 211-218, 1991.
- [31] C. Xu, J. Prince, "Gradient Vector Flow Deformable Models," Handbook of Medical Imaging, edited by Isaac Bankman, Academic Press, 2000.
- [32] H. Yoshida, S. Katsuragawa, Y. Amit, K. Doi, "Wavelet-based deformable contour and its application to detection of pulmonary nodules on chest radiographs," Proceedings of SPIE Wavelet Applications in Signal and Image Processing, vol. 3169, pp.328-336, 1997.
- [33] A. Pentland, S. Sclaroff, "Closed-form solutions for physically based modelling and recognition," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol 13, pp. 715-729, 1991.
- [34] J. Park, D. Matas, A. Young, and L. Axel, "Deformable models with parameter functions for cardiac motion analysis from tagged MRI data," IEEE Transactions on Medical Imaging, vol. 15, pp. 278-289, 1996.
- [35] T. Cootes, C. Taylor, D. Cooper, J. Graham, "Training models of shape from sets of examples," Proceedings of British Machine Vision Conference, pp. 9-18, Springer-Verlag, 1992.
- [36] T. Cootes, C. Taylor, D. Cooper, J. Graham, "Active shape models-their training and application," Journal of Computer Vision and Image Understanding, vol. 61(1), pp. 38-59, 1995.
- [37] T. Cootes, T. Taylor, "Active Appearance models," Proceedings of 5th European Conference on Computer Vision, vol. 2, Germany, 1998.
- [38] T. Cootes, C. Taylor, "Statistical Models of Appearance for Computer Vision," Imaging Science and Biomedical Engineering, U. of Manchester, Technical report, 2004.

- [39] C. Glasbey, K. Mardia, "A review of image-warping methods," *Journal of Applied Statistics*, vol. 25(2), pp. 155-172, 1998.
- [40] A. Mohamed, "Combining Statistical and Biomechanical Models for Estimation of Anatomical Deformations," PhD thesis, Johns Hopkins University, 2005.
- [41] T. Cootes, and C. Taylor, "Combining point distribution models with shape models based on finite element analysis," *Image and Vision Computing*, vol. 13(5), pp. 403-409, 1995.
- [42] Y. Wang, L. Staib, "Boundary finding with prior shape and smoothness models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22(7), pp.738-743, July 2000.
- [43] T. Cootes, and C. Taylor, "A mixture model for representing shape variation," *Image and Vision Computing*, vol. 17(8), pp. 567-574, 1999.
- [44] R. Bowden, T. Mitchell, and M. Sarhadi, "Cluster based nonlinear principle component analysis," *Electronics Letters*, vol. 33, pp. 1858-1859, 1997.
- [45] C. Davatzikos, X. Tao, D. Shen, "Hierarchical Active Shape Models, using the Wavelet Transform," *IEEE Transaction on Medical Imaging*, vol. 22(3), pp.414-423, 2003.
- [46] Z. Xue, D. Shen, B. Karacali, C. Davatzikos, "Statistical Representation and Simulation of High-Dimensional Deformations: Application to Synthesizing Brain Deformations," *Proceedings of MICCAI 2005, Lecture Notes on Computer Science*, pp. 500-508, 2005.
- [47] D. Nain, S. Haker, A. Bobick, A. Tannenbaum, "Multiscale 3D Shape Analysis using Spherical Wavelets," *Proceedings of MICCAI 2005, LNCS*, pp. 459-467, 2005.
- [48] D. Nain, S. Haker, A. Bobick, and A. Tannenbaum, "Multiscale 3-D Shape Representation and Segmentation Using Spherical Wavelets," *IEEE Transaction on Medical Imaging*, vol. 26(4), pp. 598-618, 2007.

- [49] M. Uzumcu, A. Frangi, J. Reiber, B. Lelieveldt, "Independent Component Analysis in Statistical Shape Models," *Proceedings of SPIE Medical Imaging 2003*, vol. 5032, pp. 375-383, 2003.
- [50] G. Rogez, C. Orrite-Uruñuela, J. Martínez del Rincón, "Human figure segmentation using independent component analysis," *Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA)*, pp. 300-307, Portugal, June 2005.
- [51] M. Üzümcü, A. Frangi, M. Sonka, J. Reiber, B. Lelieveldt, "ICA vs. PCA Active Appearance Models: Application to Cardiac MR Segmentation," *Proceedings of MICCAI 2003, Lecture Notes on Computer Science (LNCS)*, pp. 451-458, 2003.
- [52] J. Koikkalainen, J. Lötjönen, "Image segmentation with the combination of the PCA- and ICA- based modes of shape variation," *Proceedings of IEEE Int. Symposium on Biomedical Imaging : Macro to Nano*, pp.149-152, 2004.
- [53] K. Sjöstrand, M. Stegmann, R. Larsen, "Sparse Principal Component Analysis in Medical Shape Modeling," *Proceedings of SPIE Medical Imaging*, vol. 6144, 2006.
- [54] K. Sjöstrand, E. Rostrup, C. Ryberg, R. Larsen, C. Studholme, H. Baezner, J. Ferro, F. Fazekas, L. Pantoni, D. Inzitari, G. Waldemar, "Sparse Decomposition and Modeling of Anatomical Shape Variation," *IEEE Transaction on Medical Imaging*, vol. 26(12), pp. 1625-1635, 2007.
- [55] C. Chennubhotla, A. Jepson, "Sparse-PCA: Extracting multi-scale structure from data," *IEEE International Conference on Computer Vision 2001*, vol. 1, pp. 641-647, 2001.
- [56] M. Roberts, T. Cootes, J. Adams, "Automatic segmentation of lumbar vertebrae on digitised radiographs using linked active appearance models," *Proceedings of Medical Image Understanding and Analysis*, vol. 2, pp. 120-124, 2006.
- [57] K. Delac, M. Grgic, P. Liatsis, "Appearance-based statistical methods for face recognition," *Proceedings of the 47th International Symposium ELMAR-2005 focus on Multimedia Systems and Applications*, pp. 151-158, Croatia, 2005.

- [58] X. Xu, C. Zhang, and T. Huang, "Active Morphable Model: An Efficient Method for Face Analysis," in Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR 2004), pp. 837-842, 2004.
- [59] Y. Zhang, "Modeling of human faces with parameterized local shape morphing," 15th IEEE International Symposium on Robot and Human Interactive Communication, pp. 20-25, 2006.
- [60] M. Stegmann, K. Sjostrand, R. Larsen, "Sparse Modeling of Landmark and Texture Variability using the Orthomax Criterion," Proceedings of SPIE Medical Imaging, vol. 6144, 2006.
- [61] C. Wolstenholme, C. Taylor, "Wavelet Compression of Active Appearance Models," "Proceedings of the Second International Conference on Medical Image Computing and Computer-Assisted Intervention, p.544-554, 1999 .
- [62] M. Stegmann, S. Forchhammer, T. Cootes, "Wavelet Enhanced Appearance Modelling," Proceedings of SPIE Symposium on Medical Imaging 2004, vol. 5370, pp. 1823-1832, 2004.
- [63] R. Larsen, M. Stegmann, S. Darkner, S. Forchhammer, T. Cootes, B. Ersbøll, "Texture Enhanced Appearance Models," Computer Vision and Image Understanding, vol. 106, pp. 20-30, Elsevier, 2007.
- [64] B. Abboud, F. Davoine, and M. Dang, "Facial expression recognition and synthesis based on an appearance model," Signal Processing: Image Communication, vol. 19, pp. 723-740, 2004.
- [65] W. Seales, C. Yaun, "Improved Image Classification Using Morphing," Proceedings of 3rd Asian Conference on Computer Vision, Lecture Notes on Computer Science, vol. 1352, pp. 233-240, 1998.
- [66] S. Wolberg, G. Shin, "Polymorph: morphing among multiple images," IEEE Computer Graphics & Applications, vol.18 (1), pp.58-71, 1998.

- [67] G. Hamarneh, P. Jassi, L. Tang, "Simulation of ground truth validation data via physically- and statistically- based warps," MICCAI, vol. 1, pp. 459-467, 2008.
- [68] C. Rose and C. Taylor, "A Statistical Model of Texture for Medical Image Synthesis and Analysis," In Proceedings of Medical Image Understanding and Analysis, pages 1-4, 2003.
- [69] A. ElSafi, R. Zewail, N. Durdle, "Statistical simulation of deformations using wavelet independent component analysis," Proceedings of Visual Information Processing, SPIE Symposium on Security and Defence, vol. 6978, 2008.
- [70] Z. Xue, D. Shen, B. Karacali, J. Stern, D. Rottenberg, and C. Davatzikos, "Simulating Deformations of MR Brain Images for Validation of Atlas-based Segmentation and Registration Algorithms," Neuroimage, Vol 33(3), pp. 855-866, 2006.
- [71] I. Scott, T. Cootes, C. Taylor, "Improving Appearance Model Matching Using Local Image Structure," Proceedings of Information Processing in Medical Imaging, pp. 258-269, 2003.
- [72] B. Ginneken, A. Frangi, J. Staal, B. Romeny, M. Viergever, "Active Shape Model segmentation with optimal features," IEEE Transactions on Medical Imaging, vol. 21, pp. 924-933, 2002.
- [73] S. Y. Li, T. Zhu, and T. Z. Jiang, "Active Shape Model Segmentation Using Local Edge Structures and AdaBoost," Medical Imaging and Augmented Reality, Lecture Notes in Computer Science, vol. 3150, pp. 121-128, 2004.
- [74] M. de Bruijne, M. Nielsen, "Image Segmentation by Shape Particle Filtering," 17th International Conference on Pattern Recognition (ICPR'04), vol. 3, pp. 722-725, 2004.
- [75] D. Seghers, D. Loeckx, F. Maes, D. Vandermeulen, P. Sueten, "Minimal Shape and Intensity Cost Path Segmentation," IEEE Transactions on Medical Imaging, vol. 26(8), pp. 1115-1129, 2007.

- [76] Y. Zhu, M. Fisher and R. Zwigelaar, "Improving ASM Search Using Mixture Models for Grey-Level Profiles," Proceedings of IbPRIA-2005, Lecture Notes in Computer Science, vol. 3522, pp. 292-299, 2005.
- [77] D. Metaxas, Z. Qian, X. Huang, R. Huang, T. Chen, L. Axel, "Hybrid Deformable Models for Medical Segmentation and Registration," Proceedings of International Conference on Control, Automation, Robotics and Vision, pp. 1-6, 2006.
- [78] G. Langs, P. Peloschek, R. Donner, M. Reiter, H. Bischof, "Active Feature Models," Proceedings of International Conference on Pattern Recognition ICPR 2006, vol. 1, pp. 417-420, 2006.
- [79] M. Üzümcü, A. Frangi, J. Reiber, B. Lelieveldt, "The use of Independent Component Analysis in Statistical Shape Models," Proceedings of SPIE Medical Imaging 2003, vol 5032, pp. 375-383, 2003.
- [80] A. Mohamed, C. Davatzikos, "Shape Representation via Best Orthogonal Basis Selection," MICCAI- 2004, vol. 1, pp. 225-233, 2004.
- [81] J. Koikkalainen, J. Lotjonen, "Image Segmentation with the combination of the PCA and ICA-based modes of shape variation," IEEE International Symposium on Biomedical Imaging: Nano to Macro, vol. 1, pp. 149-152, 2004.
- [82] B. Howe, A. Gururajan, H. Sari Sarraf, L.R. Long, "Hierarchical segmentation of cervical and lumbar vertebrae using a customized generalized Hough transform and extensions to active appearance models," 6th IEEE Southwest Symposium on Image Analysis and Interpretation, pp:182-186, 2004.
- [83] L. R. Long, G. Thoma, "Identification and classification of spine vertebrae by automated methods," Proceedings of SPIE Medical Imaging 2001, vol. 4322, pp. 1478-1489, 2001.
- [84] R. Eastell, S. Cedal, H. Wahner, V. Riggs, L. Melton, "Classification of Vertebral Fractures," Journal of Bone and Mineral Research, vol. 6(3), pp. 207-215, 1991.

- [85] E. McCloskey, T. Spector, K. Eyres, E. Fern, N. O'Rourke, S. Vasikaran, and J. Kanis, "The assessment of vertebral deformity: a method for use in population studies and clinical trials," *Osteoporosis International* 3, pp. 138-147, 1993.
- [86] S. Antani, L. Long, G. Thoma, R. Stanley, "Vertebra shape classification using MLP for content-based image retrieval," *Proceedings of International Joint Conference on Neural Networks* 2003, pp. 160-165, 2003.
- [87] S. Antani, X. Xu, L. Long, G. Thoma, "Partial Shape Matching for CBIR of Spine X-ray Images," *Proceedings of SPIE Electronic Imaging, Storage and Retrieval Methods and Applications for Multimedia*, vol.5307, 2004.
- [88] S. Antani, D. Lee, L. Long, G. Thoma, "Evaluation of shape similarity measurement methods for spine X-ray image," *Journal of Visual Communication and Image Representation*, vol. 15(3), pp. 285-302, 2004.
- [89] S. Antani, L. Long, G. Thoma, "Applying Vertebral Boundary Semantics to CBIR of Digitized Spine X-ray Images," *Proceedings of IS&T/SPIE Electronic Imaging Science and Technology, Storage and Retrieval Methods and Applications for Multimedia*, vol. 5682, pp. 98-107, 2005.
- [90] P. Smyth, C. Taylor, J. Adams, "Vertebral shape: automatic measurement with active shape models," *Radiology*, vol. 211, pp. 571-578, 1999.
- [91] M. Roberts, T. Cootes, J. Adams, "Vertebral Morphometry: Semi-automatic determination of detailed vertebral shape from DXA images using Active Appearance Models," *Investigative Radiology*, vol. 41(12), pp. 849-859, 2007.
- [92] M. Roberts, T. Cootes, E. Pacheca, J. Adams, "Quantitative vertebral fracture detection on DXA images using shape and appearance models," *Academic Radiology*, vol. 14(10), pp. 1166-78, 2007.

- [93] M. de Bruijne, M. Lund, L. Tanko, P. Pattersen, and M. Nielsen, "Quantitative vertebral morphometry using neighbour-conditional shape models," *Proceedings of MICCAI-2006, Lecture Notes in Computer Science*, vol. 4190, pp. 1-8, 2006.
- [94] M. de Bruijne, P. Pattersen, L. Tanko, and M. Nielsen, "Vertebral fracture classification," *Proceedings of SPIE Medical Imaging*, vol. 6512, 2007.
- [95] B. Olshausen, D. Field, "How close are we to understanding V1?," *Neural Computing*, vol. 17, pp. 1665-99, 2005.
- [96] J. Daugman, "Gabor wavelets and statistical pattern recognition," *The Handbook of Brain Theory and Neural Networks*, 2nd edition, editor: M. Arbib, MIT Press, pp. 457-463.
- [97] Stephane Mallat, *A Wavelet tour of Signal Processing*. Academic Press, 1999.
- [98] L. Prasad, *Wavelet Analysis with Applications to Image Processing*. CPC Press, 1st edition, 1997.
- [99] M. N. Do and M. Vetterli, "Contourlets," in *Beyond Wavelets*, editor: G. V. Welland, Academic Press, 2003.
- [100] D. Gabor, "Theory of communications," *Journal IEE*, vol. 98 (3), 1946.
- [101] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by 2D visual cortical filters," *Journal of Optical Society of America (A)*, vol.2 (7), pp.1160-1169, 1985.
- [102] V. Kyrki, "Local and global feature extraction for invariant object recognition," *Ph.D. thesis, Lappeenranta University of Technology*, 2002.
- [103] Joni-Kristian Kämäräinen, "Feature Extraction Using Gabor Filters," *PhD thesis, Lappeenranta University of Technology*, 2003.

- [104] E. Simoncelli, W. Freeman, E. Adelson, D. Heeger, "Shiftable multiscale transforms," *IEEE Transactions on Information Theory*, vol.38(2), pp. 587-607, 1992.
- [105] R. Bamberger and M. Smith, "A filter bank for the directional decomposition of images: Theory and design," *IEEE Transaction on Signal Proceeding*, vol. 40 (4), pp. 2329-2342, Sep. 2003.
- [106] A. L. Da Cunha, J. Zhou, M.N. Do, "The Nonsubsampled Contourlet Transform: Theory, Design, and Applications," *IEEE Transactions on Image Processing*, vol. 15(10), pp. 3089– 3101, 2006.
- [107] G. Hamarneh, R. Abu-Gharbieh, T. Gustavsson, "Review-Active Shape Models-Part I: Modeling Shape and Gray Level Variation," *Swedish Symposium on Image Analysis*, pp. 125-128, 1998.
- [108] L. Long, S. Antani, G. Thoma, "Image informatics at a National Research Center," *Journal of Computerized Medical Imaging and Graphics*, vol. 29, pp. 171-193, 2005.
- [109] B. Olshausen, D. Field, "Sparse coding with an over-complete basis set: a strategy employed by V1?," *Vision Research*, vol. 37, pp. 3311-3325, 1997.
- [110] J. Wright, Yi Ma, J. Mairal, G. Spairo, T. Huang, S. Yan, "Sparse Representation for Computer Vision and Pattern Recognition," submitted to the *Proceedings of IEEE*, march 2009.
- [111] H. Ólafsdóttir, M. Hansen, K. Sjöstrand, T. Darvann, N. Hermann, E. Oubel, B. Ersbøll, R. Larsen, A. Frangi, P. Larsen, C. Perlyn, G. Morriss-Kay, S. Kreiborg, "Sparse Statistical Deformation Model for the Analysis of Craniofacial Malformation in the Crouzon Mouse," *Scandinavian Conference on Image Analysis 2007, Lecture Notes in Computer Science*, vol. 4522, pp. 112-121, 2007.
- [112] M. Wickerhauser, *Adapted wavelet analysis from Theory to Software*. IEEE press, Piscataway, New Jersey, 1998.

- [113] J. Shi, J. Malik, "Normalized cuts and image segmentation," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 22(8), pp. 888-905, 2000.
- [114] Y. Jin, E. Angelini and A. Laine, "Wavelets in Medical Image Processing: Denoising, Segmentation, and Registration," *Handbook of Medical Image Analysis: Advanced Segmentation and Registration Models*, edited by Jasjit Suri, David L. Wilson and Swamy Laximinarayan. Kluwer Academic Publishers, New York NY, 2004.
- [115] R. Coifman, M. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Transactions on Information Theory*, vol. 38(2), pp. 713-718, 1992.
- [116] A. Hyvarinen, J. Karhunen, E. Oja, "Independent Component Analysis," John Wiley & Sons, 2001.
- [117] A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE Transactions on Neural Networks*, vol. 10(3), pp.626-634, 1999.
- [118] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer-Verlag, 2001.
- [119] A. Webb, *Statistical Pattern Recognition*. 2nd edition, 2002.
- [120] M. Cherukuri, R. Stanley, R. Long, S. Antani, G. Thoma, "Anterior osteophyte discrimination in lumbar vertebrae using size-invariant feature," *Journal of Computerized Imaging and Graphics*, vol. 28, pp. 99-108, 2004.
- [121] R. Stanley, S. Antani, R. Long, G. Thoma, K. Gupta, M. Das, "Size-invariant descriptors for detecting regions for abnormal growth in cervical vertebrae," *Journal of Computerized Medical Imaging and Graphics*, vol. 32, pp. 44-52, 2008.
- [122] A. Suinesiaputra, M. Üzümcü, A. Frangi, T. Kaandorp, J. Reiber, B. Lelieveldt, "Detecting Regional Abnormal Cardiac Contraction in Short-Axis MR Images Using Independent Component Analysis," *MICCAI 2004*, vol. 1, pp. 737-744, 2004.

- [123] T. Nicholas, A. Holmes, "Nonparametric permutation tests for functional neuroimaging: A primer with examples," *Human Brain mapping*, vol. 15, pp. 1-25, 2002.
- [124] F. Heijden, R. Duin, D. Ridder, *Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB*. Wiley, 2004.
- [125] T. Deselaers, T. Deserno, H. Müller, "Automatic medical image annotation in ImageCLEF 2007: Overview, Results, Discussion," *Pattern Recognition Letters*, vol. 28, pp.1988-1995, 2008.
- [126] C. Schmid, R. Mohr and C. Bauckhage, "Evaluation of Interest Point Detectors," *International Journal of Computer Vision*, vol. 37(2), pp.151-172, 2000.
- [127] C. Harris, M. Stephens, "A combined corner and edge detector," *Proceedings of Alvey Vision Conference*, 1988.
- [128] D. Lowe, "Object Recognition from Local Scale-Invariant Features," *Proceedings of 7th IEEE International Conference on Computer Vision*, vol. 2, pp. 1150-1157, 1999.
- [129] Q. Tian, N. Sebe, M. S. Lew, E. Louprias, and T. S. Huang, "Image retrieval using wavelet-based salient points," *Journal of Electronic Imaging*, vol. 10(4), pp. 835-849, 2001.
- [130] E. Louprias, N. Sebe, "Wavelet-based Salient Points: Applications to Image Retrieval Using Color and Texture Features," *Proceedings of 4th Int. Conf. on Visual Information Systems*, pp. 223-232, 2000.
- [131] L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 20(11), pp. 1254-1259, 1998.
- [132] C. Wu, Q. Wang, "A Novel Approach for Interest Point Detection Based on Phase Congruency," *Proceedings of TENCON 2005*, pp. 1-6, 2005.

- [133] X. Gao, F. Sattar, R. Venkateswarlu, "Multiscale Corner Detection of Gray Level Images Based on Log-Gabor Wavelet Transform," *IEEE Transactions on Circuits Systems and Video Technology*, vol. 17(7), pp. 868-875, 2007.
- [134] P. Kovesi, "Phase Congruency Detects Corners and Edges," *Australian Pattern Recognition Society Conference: DICTA 2003*, pp. 309-318, 2003.
- [135] M. Nixon, A. Aguado. *Feature Extraction and Image Processing*. 2nd edition, Elsevier, 2008.
- [136] F. Bookstein, "Principal Warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11(6), pp.567-585, 1989.
- [137] R. Sprengel, K. Rohr, H. Stiehl, "Thin-Plate Spline Approximation for Image Registration," *Proceedings of 18th Annual International Conference of IEEE Engineering in Medicine and Biology*, pp. 1190–1191, 1996.
- [138] A. X. Falcao, J. K. Udupa, S. Samarasekera, and S. Sharma, "User-Steered Image Segmentation Paradigms: LiveWire and Live Lange," *Graphical Models and Image Processing*, pp-233-260, 1998.
- [139] E. Mortensen, W. Barrett, "Intelligent Scissors for Image Composition, "Computer Graphics (SIGGRAPH'95), pp. 191-198, 1995.
- [140] L. O'Donnell, C. Westin, W. Grimson, J. Ruiz-Alzola, M. Shenton, R. Kikinis, "Phase-based user-steered image segmentation," *International conference of Medical Image Computing and Computer Assisted Intervention (MICCAI-2001)*, vol. 4, pp. 1022-1030, 2001.
- [141] S. Seetharaman, R. Stanley, S. De, S. Antani, R. Long, G. Thoma, "Livewire Segmentation Tool for Osteophyte Detection in Lumbar Spine X-Ray Images," 1st *International Conference on Emerging Trends in Engineering and Technology*, pp. 715-720, 2008.

- [142] T.Cormen, C. Leiserson, R. Rivest, and C.Stein. Introduction to Algorithms, 2nd edition, MIT Press and McGraw-Hill, 2001. Section 24.3: Dijkstra's algorithm, pp.595–601.
- [143] J. Kalpathy-Cramer, W. Hersh, “Medical Image Retrieval and Automatic Annotation: OHSU at ImageCLEF 2007,” Lecture notes on Computer Science, pp. 623-630, 2008.
- [144] T. Deselaers, D. Keysers and H. Ney, “Features for image retrieval: an experimental comparison,” Journal of Information Retrieval, vol.11, pp. 77-107, 2008.
- [145] W. Liu, Q. Y. Tong, “Medical Image Retrieval using Salient Point Detector,” 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp:6352-6355, 2006.
- [146] R. Maree, P. Geurts, J. Piater, L. Wehenkel, “Random sub windows for robust image classification,” In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition-2005, pp.34-40, 2005.
- [147] T. Tommasi, F. Orabona, B. Caputo, “CLEF2007 Image Annotation Task: an SVM-based Cue Integration Approach,” In Working Notes of the 2007 CLEF Workshop, Budapest, Hungary, 2007.
- [148] P. Howarth, A. Yavlinsky, D. Heesch, S. Rüger, “Medical Image Retrieval Using Texture, Locality, and Colour,” CLEF 2004, pp. 740-749, 2004.
- [149] C. Ngo, T. Pong, R. Chin, “Exploiting Image Indexing Techniques in DCT Domain,” Proceedings of IAPR International Workshop on Multimedia Information Analysis and Retrieval, Lecture Notes in Computer Science, vol. 1464, pp:195-206, 1998.
- [150] A. Selvarani, S. Annadurai, “Medical Image Retrieval by Combining Low Level Features and DICOM Features,” Proceedings of International Conference on Computational Intelligence and Multimedia Applications, vol. 1, pp. 587-589, 2007.

- [151] <http://www.robots.ox.ac.uk/~vgg/data/data-aff.html>.
- [152] K. Mikolajczyk, C. Schmid, "A Performance Evaluation of Local Descriptors," Proceedings of Conference of Computer Vision and Pattern Recognition, pp. 257-264, 2003.
- [153] W. Freeman, E. Adelson, "The design and use of Steerable filters," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.13 (9), pp. 891-906, 1991.
- [154] G. Zamora, H. Sari Sarraf, L. Long, "Hierarchical Segmentation of Vertebrae from X-ray images," Proceedings of SPIE Medical Imaging: Image Processing, vol. 5032, pp. 631-42, 2003.
- [155] C. Rao, S. Kumar, B. Chatterji, "Content Based Image Retrieval using Contourlet Transform," International Journal on Graphics, Vision and Image Processing, vol. 7(3), pp. 9-15, 2007.
- [156] M. Nordstom, M. Larsen, J. Sierakowski, and M. Stegmann, "The IMM Face Database: An Annotated Dataset of 240 Face Images, Informatics and Mathematical Modelling," Technical Univ. of Denmark, DTU, Denmark, 2004.
- [157] P. Burt, E. Adelson, "The Laplacian pyramid as a compact image coding, IEEE Trans. Comm. Com, Vol. 31:pp 532--540, 1983.

APPENDIX A: Aligning Shapes

In this appendix, we describe the procedure to align a set of shapes.

Aligning two shapes:

Given two shapes $\mathcal{X}_1$ and $\mathcal{X}_2$, we want to estimate the transformation parameters that best align the two shapes with respect to a weighted sum of the square distance. The transformation parameters are rotation θ , scale s , and translation $\mathbf{t} = (t_x, t_y)$.

The weighted mean square error is defined as

$$E = [\mathbf{x}_1 - M(s, \theta)(\mathbf{x}_1) - \mathbf{t}]^T W [\mathbf{x}_1 - M(s, \theta)(\mathbf{x}_1) - \mathbf{t}], \quad (\text{A.1})$$

where :

$$M(s, \theta) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_x x - a_y y \\ a_y x + a_x y \end{pmatrix},$$

$$a_x = s \cos(\theta) \quad \text{and} \quad a_y = s \sin(\theta).$$

The least square solution is obtained by differentiating equation A.1 with respect to each of the parameters. This procedure produces a set of equations that can be written in matrix form as follows:

$$\begin{pmatrix} X_2 & -Y_2 & W & 0 \\ Y_2 & X_2 & 0 & W \\ Z & 0 & X_2 & Y_2 \\ 0 & Z & -Y_2 & X_2 \end{pmatrix} \begin{pmatrix} a_x \\ a_y \\ t_x \\ t_y \end{pmatrix} = \begin{pmatrix} X_1 \\ Y_1 \\ C_1 \\ C_2 \end{pmatrix}, \quad (\text{A.2})$$

where:

$$X_i = \sum_{j=0}^{N-1} w_k x_{ij}, \quad Y_i = \sum_{j=0}^{N-1} w_k y_{ij}, \quad Z = \sum_{j=0}^{N-1} w_k (x_{2j}^2 + y_{2j}^2),$$

$$W = \sum_{k=0}^{N-1} w_k, \quad C_1 = \sum_{j=0}^{N-1} w_j (x_{1j} x_{2j} + y_{1j} y_{2j}), \quad C_2 = \sum_{j=0}^{N-1} w_j (y_{1j} x_{2j} - x_{1j} y_{2j}).$$

Equation A.2 is then solved by using standard matrix methods.

Aligning multiple shapes:

A set of shapes is then aligned in the following iterative approach:

1. Select a shape as an estimate of the mean shape.
2. Align all remaining shapes to the estimated mean shape.
3. Re-estimate the mean shape from the aligned shapes.
4. Go to step 2 until convergence.

APPENDIX B: Parameters Update for ASM Search

This appendix describes the procedure for updating the shape parameters in the Active Shape Model (ASM) image search. The shape parameters in ASM include

The pose parameters: (scale s , rotation θ , translation $\mathbf{t}$),

The PCA-model parameter: (weights of eigenvectors $\mathbf{b}$).

We start with an initial shape estimation x_i with the initial pose parameters ($s_i, \theta_i, \mathbf{t}_i$),

and the initial PCA-model parameters set to zero ($x_m = \bar{x} + \phi\mathbf{b}, \mathbf{b} = \mathbf{0}$):

$$x_i = M(s_i, \theta_i)[\bar{x}] + \mathbf{t}_i. \quad (\text{B.2})$$

Let dx_i be the suggested movement of the current shape estimate. Then we want to update the model parameters so that the current shape estimate best fits the suggested movement in the image:

$$x_i \xrightarrow{(ds, (1+d\theta), dt, d\mathbf{b})} x_i + dx_i. \quad (\text{B.3})$$

Find new pose parameters:

The current estimate of the pose parameters is first updated to best fit the image. This step is done by finding the displacements to the pose parameters ($ds, (1+d\theta), dt$) by using the method in Appendix A:

$$x_i \xrightarrow{(ds, (1+d\theta), dt)} x_i + dx_i \quad (\text{B.4})$$

$$x_i = M(s_i, \theta_i)[\bar{x}] + \mathbf{t}_i \xrightarrow{(ds, (1+d\theta), dt)} x_i + dx_i. \quad (\text{B.5})$$

Find new PCA-model parameters:

Knowing the pose displacements, ($ds, (1+d\theta), dt$), we solve for the PCA-model parameters (dx_m):

$$M(s(1+ds), (\theta+d\theta))[\bar{x} + dx_m] + (\mathbf{t} + dt) = x + dx \quad (\text{B.6})$$

$$M(s_i(1+ds), (\theta_i + d\theta))[\bar{x} + dx_m] = M(s_i, \theta_i)[\bar{x}] + dx_m - d\mathbf{t} . \quad (\text{B.7})$$

Since $M^{-1}(s, \theta)[.] = M(s^{-1}, -\theta)[.]$, we have

$$\bar{x} + dx_m = M((s_i(1+ds))^{-1}, -(\theta_i + d\theta)).[M(s_i, \theta_i)[\bar{x}] + dx - d\mathbf{t}] . \quad (\text{B.8})$$

Then

$$dx_m = M((s_i(1+ds))^{-1}, -(\theta_i + d\theta)).[M(s_i, \theta_i)[\bar{x}] + dx - d\mathbf{t}] - \bar{x} \quad (\text{B.9})$$

$$\bar{x} + dx_m = \bar{x} + \phi d\mathbf{b} . \quad (\text{B.10})$$

Thus, $d\mathbf{b}$ is given by: $d\mathbf{b} = \phi^T dx_m$.

During the image search, the parameters are updated in an iterative manner as

$$\mathbf{t}_{i+1} \rightarrow \mathbf{t}_i + W_t d\mathbf{t}_i, \quad (\text{B.11})$$

$$\theta_{i+1} \rightarrow \theta_i + W_\theta d\theta_i, \quad (\text{B.12})$$

$$s_{i+1} \rightarrow s_i(1 + W_s ds_i), \quad (\text{B.13})$$

$$\mathbf{b}_{i+1} \rightarrow \mathbf{b}_i + W_b d\mathbf{b}_i. \quad (\text{B.14})$$

C12088